

AD _____
(Leave blank)

Award Number: W81XWH-10-1-0989

TITLE: Largazole as a Novel and Selective Anti-Breast Cancer Agent

PRINCIPAL INVESTIGATOR:

Xuedong Liu, Ph.D.

CONTRACTING ORGANIZATION:

University of Colorado at Boulder,
Dept. of Chemistry and Biochemistry
UCB215, Boulder, CO 80309

REPORT DATE:

October 2013

TYPE OF REPORT:

Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

✓ Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.		
1. REPORT DATE (DD-MM-YYYY) October 2013	2. REPORT TYPE Final Addendum	3. DATES COVERED (From - To) 17September2012-16Septemkber2013
4. TITLE AND SUBTITLE Largazole as a Novel and Selective Anti-Breast Cancer Agent		5a. CONTRACT NUMBER
		5b. GRANT NUMBER BC095674
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Xuedong Liu		5d. PROJECT NUMBER W81XWH-10-1-0989
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Colorado at Boulder, Dept. of Chemistry and Biochemistry UCB215, Boulder, CO 80309		8. PERFORMING ORGANIZATION REPORT NUMBER REPORT NUMBER 1536128
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Contracts and Grants University of Colorado at Boulder 576 UCB Boulder, CO 80309		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.		
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Defense position, policy or decision, unless so designated by other documentation.		
14. ABSTRACT Breast cancer is the second most common type of cancer in the world and second most common cause of deaths in US. There is great a demand for new, small molecule, drugs that can selectively eliminate breast cancer cells. Many natural compounds have anti-tumor activities (e.g. Taxol®). Recently we achieved a total synthesis of Largazole and demonstrated that this natural compound has remarkable selectivity toward breast cancer cells. We also find that Largazole can block two cellular activities frequently associated with aggressive tumor cells. In this proposal, we will test whether dual inhibition of two oncogenic pathways may be the reason why Largazole is highly selective against tumor cells but not normal cells. We will develop more potent and selective small molecules to validate the concept that dual specificity inhibitors are better anti-cancer drugs. Our studies are expected to provide novel ideas for designing more effective therapeutics for breast cancer treatment.		

15. SUBJECT TERMS Largazole, ubiquitin E1, Histone deacetylase, inhibitor, breast cancer,					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Xuedong Liu
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED	UL		19b. TELEPHONE NUMBER <i>(include area code)</i> 3037356161

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

	<u>Page</u>
Introduction.....	5
Body.....	6-12
Key Research Accomplishments.....	12
Reportable Outcomes.....	12
Conclusion.....	13
References.....	14
Appendices.....	15

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Final report:

1. INTRODUCTION

Histone deacetylases (HDACs) and histone acetylases (HATs) are key players in regulating transcription and histone homeostasis(8). Transcription of tumor suppressor proteins is frequently silenced in tumor cells due the hyper- or aberrant activity of HDACs. Accordingly inhibiting histone deacetylation may re-activate inappropriately silenced genes and may be able to “reverse” malignant changes(2). Inhibitors of histone deacetylase enzymes (HDACi) have recently attracted substantial attention as potential anti-cancer drugs. The selective degradation of many regulatory proteins in eukaryotic cells is mediated by the ubiquitin system(5). Proteins targeted for degradation are usually covalently ligated to a polyubiquitin chain and subsequently eliminated by the 26S proteasome. Ubiquitination of proteins is carried out by a multi-enzyme complex consisting of E1 (ubiquitin activating enzyme), E2 (ubiquitin conjugating enzyme) and E3 (ubiquitin ligase) (5). The final product of this reaction is formation of a polyubiquitinated protein with attachment of an ubiquitination through an isopeptide bond to an epsilon-amino group of certain Lys residues in the interior of the substrate. There is only one ubiquitin E1 enzyme, more than fifty ubiquitin E2 and perhaps thousands of E3 enzymes in human genome. E3 often controls the specificity and timing of substrate ubiquitination (5). Both HDAC inhibitors and ubiquitin-proteasome inhibitors have found applications in treating specific type of human tumors. However, either type of inhibitor alone does not appear to exhibit a broad spectrum of inhibition in treating a variety of human cancers. These observations have prompted investigations using a combination of both types of inhibitors in anti-tumor studies. It was found that bortezomib killed multiple myeloma cells more efficiently when combined with histone deacetylase inhibitors(11). Thus, administering two inhibitors simultaneously targeting both pathways could be a feasible therapeutic strategy for cancer treatment.

The goal of this study is to test that dual-specificity small molecules capable of targeting two or more aberrant signaling pathways associated with human cancers will be more efficacious in suppressing human tumors. We found that Largazole, a cyclic depsipeptide natural product isolated from the marine cyanobacterium by Luesch and coworkers (15), has remarkable potency in selectively inhibiting the proliferation of breast cancer cells without significant effects on normal breast mammary epithelial cells. We found Largazole also inhibits ubiquitin E1. The inhibitory activity of these small molecules on ubiquitin conjugation has been traced to their inhibition of the ubiquitin E1 enzyme. To further dissect the mechanism of E1 inhibition, we analyzed the effects of these inhibitors on each of the two steps of E1 activation. We show that Largazole and its derivatives specifically inhibit the adenylation step of the E1 reaction while having no effect on thioester formation. E1 inhibition appears to be specific to human E1 as Largazole ketone fails to inhibit the activation of Uba1p, a homolog of E1 in *Schizosaccharomyces pombe*. Moreover, Largazole analogs do not significantly inhibit SUMO E1. Thus, Largazole and select analogues are novel classes of ubiquitin E1 inhibitors and valuable tools for studying ubiquitination in vitro. This class of compounds could be further developed and potentially be a useful tool in cells. Using breast cancer 3D culture model and xenograft models, we show that Largazole and our newly made Paragazole have potent antitumor activity toward triple negative breast cancer. Our ultimate

goal is to demonstrate that dual targeting of both pathways is the underlying mechanism for the potency and selectivity of Largazole for breast cancer cells.

2. BODY---Studies and Results

Three specific aims were proposed in the original application. We describe our progress in the context of approved SOW.

Aim 1. Synthesize derivatives of largazole with potentially improved molecular properties and improved selectivity for transformed vs non-transformed cells.

Task 1.1 Synthesize initial round of largazole analogs for SAR testing (**Phillips, Months 1-6**) (**Completed**)

Task 1.2 After initial SAR data is obtained, design and synthesize a second round of focused largazole analogs (**Phillips, Months 7-12**) (**Completed**)

Task 1.3 Write and submit manuscripts describing the initial phase of SAR studies as well as annual report to CDMRP (**Liu and Phillips, Month 12**) (**Completed**)

Task 1.4 Synthesize larger quantities of selected largazole derivatives for detailed testing (**Phillips, Months 14-16**) (**ongoing**)

Task 1.5 Synthesize largazole derivatives targeted to specific cancer cells e.g. folic acid derivatives (**Phillips, Months 17-23**) (**ongoing**)

Task 1.6 Testing the activity of largazole derivatives using HDAC1 enzymatic assay, p27 ubiquitination and E1 thiolester assays (**Liu, Completed**).

Results

Largazole stabilizes GFP-p27 expression in Kip16 cells. A hallmark of many advanced cancers is an excessive degradation of the cyclin-dependent kinase inhibitor p27, which is directed by SCF^{Skp2}-mediated ubiquitination. Hence, stabilization of p27 degradation represents a rational approach in cancer therapeutics. To identify small molecule inhibitors that can stabilize p27Kip1, we performed a screen of ~3000 compounds from NCI DTP diversity set along with several natural products in our collection. For the cell based screen, we generated a mink lung epithelial cell line (Kip16) stably expressing p27 that was cloned in frame with green fluorescent protein (GFP). The resulting N-terminal GFP-p27 fusion, detectable by fluorescence microscopy, was used to determine the levels of p27 expression upon treatment of cells with the compound libraries in 96-well format. Much to our surprise, the most potent hit that emerged from this screen was the natural compound Largazole (Figure 1), which was first described by Luesch and coworkers (15) and subsequently synthesized in several laboratories including ours (1, 3, 10, 14, 15, 18, 19). Largazole induced a robust and highly uniform upregulation of GFP-p27 at concentrations as low as 1 nM (Figure 2a) as compared to the expression levels after treatment with the proteasome inhibitor MG132. We did not observe an increase in GFP-p27 expression upon treatment with the vehicle control DMSO. This result suggests that Largazole can stabilize GFP-p27 expression in cultured cells.

Largazole and select analogues inhibit the *in vitro* ubiquitination of p27 and Trf1. Initial investigation into the mechanism of Largazole indicated that the compound stabilized the expression of p27 in cells. Since the concentration of cyclin-dependent kinase inhibitor p27 is mainly regulated at the protein level by increased polyubiquitination and subsequent proteasomal degradation, we hypothesized that Largazole and synthetic analogues stabilize p27 by inhibiting the ubiquitination pathway (9, 13). One of the downsides of cell based assays is that the effects observed may be attributed to the influence of multiple pathways. For example, inhibiting the proteasome, elevating transcription of GFP-p27, or inhibiting Cdk activity can also lead to an increase in p27 expression. To tease out the mechanism and action of Largazole on p27

stabilization, we decided to test the effect of Largazole on p27 ubiquitination in a fully reconstituted system *in vitro* (16, 17). To test if Largazole affects p27 ubiquitination *in vitro*, we incubated Largazole (L) with p27, ubiquitin E1, E2, SCF^{Skp2}, and Cks1. As shown in Figure 1, adding Largazole significantly reduced polyubiquitinated p27, suggesting that Largazole can block p27 ubiquitination. Since Largazole is known to be a histone deacetylase inhibitor and has a thioester moiety that links an aliphatic chain to the core, we decided to test whether inhibition of p27 degradation can be linked to its histone deacetylase inhibitory activity. The structure-activity relationship for Largazole is relatively well understood (12). Therefore we next tested a series of Largazole analogues to get a preliminary structure-activity relationship on p27 ubiquitination. To investigate this, Largazole ester (E), Largazole ketone (K), Largazole macrocycle (M), and *seco*-Largazole (S) were tested in an *in vitro* p27 ubiquitination assay (Figure 2b). We also added the HDAC inhibitor Trichostatin A (TSA), the structure of which can be found in Figure 1, to the assay to determine whether or not other HDAC inhibitors affect p27 ubiquitination. We observed that Largazole (L), Largazole ketone (K), and Largazole ester (E) inhibited the ligation of ubiquitin onto p27; however, the M and S analogues and TSA failed to inhibit the ubiquitination of p27 (Figure 1). This result suggests both the macrocycle and aliphatic chain are required for ubiquitin E1 inhibition. Furthermore, the result also suggests that the thioester moiety of Largazole is not required for inhibition, because the ketone and ester analogues were equally potent in blocking p27 ubiquitination. In addition, E1 inhibition is unrelated to HDAC inhibitor activity of Largazole as both ketone and ester failed to inhibit HDAC and TSA, a known HDAC inhibitor, does not block p27 ubiquitination *in vitro*. Prior to ubiquitination, p27 is phosphorylated by the Cdk2-CyclinE complex. We carried out an *in vitro* p27 phosphorylation assay in the presence of either DMSO or Largazole in order to test whether or not the decrease in p27 ubiquitination was due to the inhibition of the Cdk2-CyclinE complex. We observed that Largazole does not inhibit the phosphorylation of p27 (Figure 2b); therefore, the inhibition of p27-ubiquitin conjugation is due to an inhibition of the ubiquitination process rather than phosphorylation step.

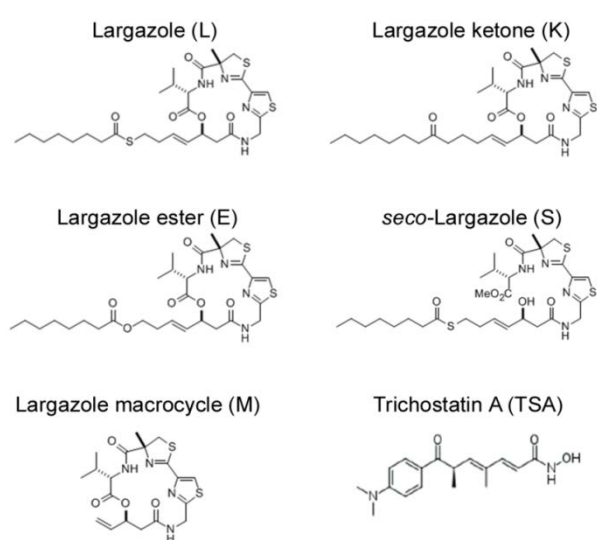


Figure 1. Chemical structures of Largazole, synthetic analogues, and Trichostatin A synthesized.

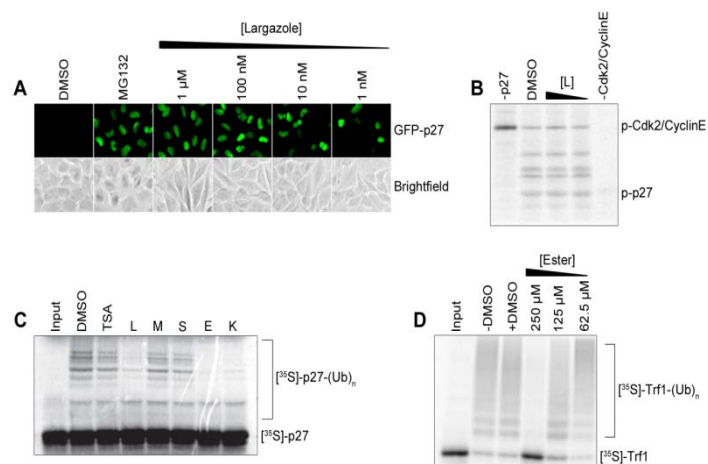


Figure 2. Largazole stabilizes p27 expression in Kip16 cells and inhibits p27 ubiquitination *in vitro* but not phosphorylation of Cdk2-CyclinE.

Even though there is a limited number of proteins in the reconstituted p27 ubiquitination system *in vitro*, tracing the real target of Largazole is still quite challenging. Fortunately, we have previously established another reconstituted *in vitro* ubiquitination assay of Trf1 with SCF^{Fbx4} (20). There are a few overlapping components between these two assays. The effect of Largazole on Trf1 ubiquitination should offer some insight as to where Largazole might target. To study the specificity of Largazole, we added Largazole ester to an *in vitro* Trf1 assay and found that Largazole ester inhibited the ligation of ubiquitin onto Trf1 in a dose-dependent fashion. Since Trf1 and p27 require different E2 ubiquitin-conjugating enzymes and different E3 ubiquitin-ligating recognition subunits in order to carry out each ubiquitination, we

hypothesized that Largazole and select synthetic analogues inhibit a step common to both ubiquitination pathways.

Largazole ketone inhibits ubiquitin E1 activation. In vertebrates, there exists only one known ubiquitin-activating E1 enzyme, UBA1. Since both p27 and Trf1 can be ubiquitinated in the presence of UBA1, we hypothesized that the inhibitory activity of Largazole is due to the deactivation of E1. To test this hypothesis, we incubated Largazole and Largazole ketone with recombinant E1 prior to carrying out an *in vitro* thioester assay we described previously (7). The presence of a fluorescence signal in the thioester assay suggests the formation of E1-ubiquitin adducts. The dose dependent decrease in fluorescence indicates that Largazole and Largazole inhibit the formation of E1-ubiquitin adducts (Fig 3AC). The dose-response curves generated from Figure 3BD suggest an IC_{50} of approximately 29 μ M and 25 μ M, respectively.

Activated ubiquitin is normally transferred to ubiquitin conjugating enzymes (E2). If E1 activity is inhibited, we expect to see that defects in E1 activation should impair the attachment of ubiquitin onto Cdc34 (E2). To further validate E1 inhibition, we included Cdc34, the E2 enzyme required for p27 ubiquitination, in the E1 reaction mixture. As shown in Figure 3EF, in the presence of ATP, fluorescent ubiquitin is transferred to Cdc34 indicated by the presence of a fluorescent Cdc34 band on the gel. Upon incubation with E2, Largazole or Largazole ester reduce the amount of ubiquitin molecules that are transferred from E1 to E2 in a dose-dependent fashion. This result is consistent with the notion that Largazole or Largazole ester inhibit E1 activity.

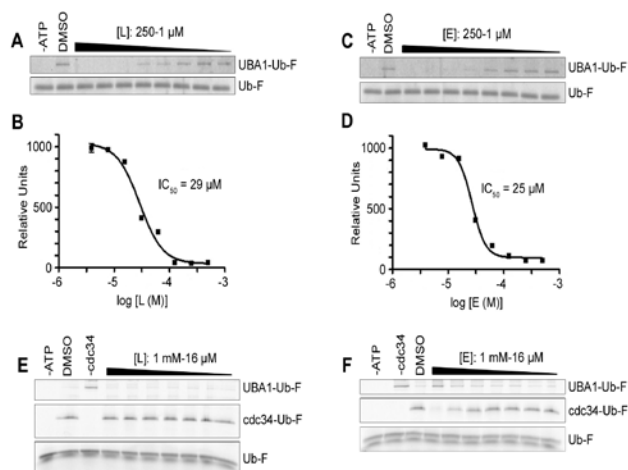


Figure 3. Largazole (L) and largazole ester (E) inhibit ubiquitin E1 in a dose dependent manner in vitro.

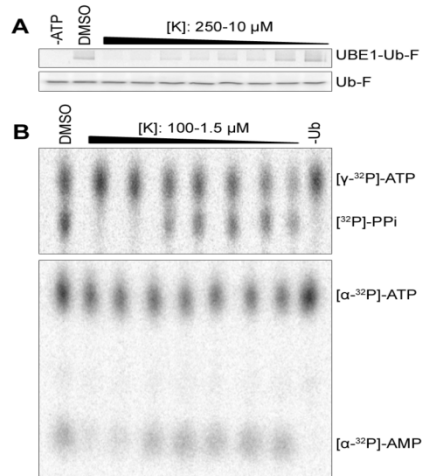


Figure 4. Largazole ketone inhibits the adenylation of the E1 ubiquitin-activating enzyme.

A potential caveat of the above experiment is that if Largazole or Largazole ester also blocks the transfer of ubiquitin from E1 to E2 we would have seen the same result. To rule out this possibility, we first produced ubiquitin charged E1 by incubating ATP and fluorescent ubiquitin for 30 min at room temperature followed by the addition of Cdc34, which was also contained with Largazole or Largazole ester. If either compound block ubiquitin transfer from E1 to E2, we would observe a significant decrease in Cdc34 fluorescence signal regardless of the order of compound addition. On the other hand, we should see the opposite results. As shown in Figure 3GH, Cdc34 is fully conjugated with fluorescence ubiquitin when Largazole or Largazole ester was added after generating fluorescent ubiquitin-E1. This result suggests that Largazole or Largazole ester neither blocks the transfer of activated ubiquitin from E1 to E2 nor promotes hydrolysis of ubiquitin thioester.

Largazole ketone inhibits the adenylation step of E1 activation. E1 forms an ubiquitin-adenylate intermediate during the course of its catalytic cycle (4). Thus the mechanism of ubiquitin E1 activation can be studied by assaying ATP:PPi and AMP:ATP exchanges (4). Production of AMP in the $[\alpha\text{-}^{32}\text{P}]\text{-AMP}:[\alpha\text{-}^{32}\text{P}]\text{-ATP}$ exchange assay guarantees that a thioester bond is formed between E1 and ubiquitin, while the release of PPi, measured by the $[\text{P}^{32}\text{P}]\text{-PPi}:[gamma\text{-}^{32}\text{P}]\text{-ATP}$ exchange assay, signals the formation of ubiquitin

adenylate. To further dissect the mechanism of Largazole inhibition, two nucleotide exchange assays were carried out in the presence of Largazole derivatives. For these experiments we used Largazole ketone, which is similar to Largazole and Largazole ester. From the results shown in Figure 4, it is evident that the first two concentrations of Largazole ketone (100 and 50 μM) inhibit ubiquitination of E1 similarly and were also inhibitory in both types of exchange assays. The lack of a $[\text{}^{32}\text{P}]\text{-PPi}$ signal suggests that the adenylation step did not occur; consequently, ubiquitin could not be transferred to the active site cysteine to trigger the release of AMP. Both steps of the E1-catalyzed reactions can be measured by the AMP:ATP exchange assay. The lack of an $[\alpha\text{-}^{32}\text{P}]\text{-AMP}$ signal further suggests that the adenylation step is inhibited by Largazole ketone. Thus Largazole or Largazole derivatives act on the first step of ubiquitin activation pathway by blocking the formation of ubiquitin-adenylate.

Selectivity of Largazole ketone against SUMO E1 and Uba1p.

In addition to ubiquitin, there exist several ubiquitin-like proteins that covalently modify other proteins. All of the ubiquitin-like proteins have activation pathways similar to ubiquitin (6). In order to study the specificity of Largazole to the ubiquitin pathway, we incubated Largazole ketone with SUMO-activating E1 enzyme prior to carrying out a thioester assay. From the results in Figure 5b, we found that Largazole ketone is ineffective in inhibiting the formation of E1-SUMO adducts. From the dose-response curve generated from the SUMO E1 fluorescence results, the IC_{50} is approximately 450 μM as opposed to 10 μM for ubiquitin E. Thus Largazole is relatively selective in perturbing ubiquitin E1 activation.

Ubiquitin and the ubiquitin E1 enzyme are highly conserved among eukaryotes (6). Sequence analysis shows a 45% homology between the human ubiquitin-activating enzyme E1 (UBA1) and *S. pombe* E1 (ptr3/Uba1p) at the amino acid sequence level. To test whether Largazole ketone inhibits the *S. pombe* E1, we carried out a thioester assay using Largazole ketone and the ubiquitin E1 homologue in *S. pombe*, Uba1p. The results in Figure 5a suggest that Largazole ketone fails to inhibit the formation of E1-ubiquitin adducts at concentrations less than 1 mM. Taken together, these results suggest that Largazole and its derivative are highly selective in inhibiting the ubiquitin E1 enzyme.

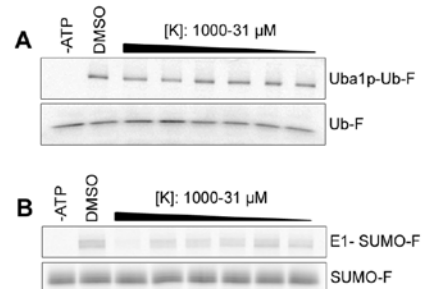


Figure 5. Investigation into the selectivity of Largazole ketone (K). A, Largazole ketone (K) fails to inhibit the ligation of ubiquitin onto Uba1p, a homologue of UBA1 from *S. pombe*. Formation of Uba1p-ubiquitin adducts was determined by thioester assay utilizing fluorescein-ubiquitin.

The results obtained in Aim 1 have been published in PLoS ONE Journal earlier this year.

Aim 2. To determine whether dual inhibition of both HDAC and ubiquitin conjugation is responsible for the selectivity of largazole against breast cancer cells and determine which HDAC isoforms render breast epithelial cells sensitive to largazole.

We have started the effort of cloning of all HDAC enzymes. So far we have cloned HDAC1, 2, 3, 4, 6, 8 and 11. We will continue to procure clones from ATCC to complete the HDAC enzyme set cloning project. The TGI assay on MDA-MB231 and a battery of breast cancer cell lines have been ongoing.

- Task 2.1 Cloning of 17 HDAC enzymes into lentiviral vector (**Liu, Months 1-4**) (**Completed**).
- Task 2.2 Construct human mammary epithelial cell lines (HME) expressing each individual HDAC enzyme (**Liu, Months 5-16**) (**Partially completed**).
- Task 2.3 Perform cell based growth inhibition assays using MDA-MB231 and HME cells using largazole analogs generated in Aim 1 (**Liu, Months 6-18**) (**ongoing**).
- Task 2.4 Measure cell permeability of largazole analogs using parallel artificial membrane permeability assay (**Liu, Months 6-18**) (**completed**).

- Task 2.5 Measuring the largazole sensitivity of HME cell lines expressing HDAC enzymes (**Liu, Months 12-18) (Completed).**
- Task 2.6 Perform siRNA and shRNA experiments for informative HDAC enzymes in HME and MDA-MB231 cells (**Liu, Months 19-24) (completed).**
- Task 2.7 Write the manuscript describing dual targeting activity of largazole (**Liu and Phillips, Month 12, completed).**

We tested a panel of 18 breast cancer cell lines from a heterogeneous group of breast cancer cell lines. In our initial exploratory studies, we made an interesting observation that triple negative breast cancer (TNBC) cell lines are particularly sensitive to Paragazole inhibition. Since TNBC are associated with a shorter median time to relapse and death and significant unmet medical need due to the fact that these cancers do not respond to endocrine therapy or other available targeted agents, we decide to focus on testing TNBC cell lines instead of broader spectrum of breast cancer cell lines. We used a panel of 19 breast cancer cell lines to assess the proliferative response to increasing concentrations of Paragazole using an SRB assay. As shown in Figure 6, a majority of these cell lines are quite sensitive (IC_{50} =500 nM) to Paragazole treatment although some are more sensitive than others. Several TNBC lines are inhibited at subnanomolar concentrations of Paragazole. These preliminary data suggests that Paragazole might be a novel agent to combat TNBC.

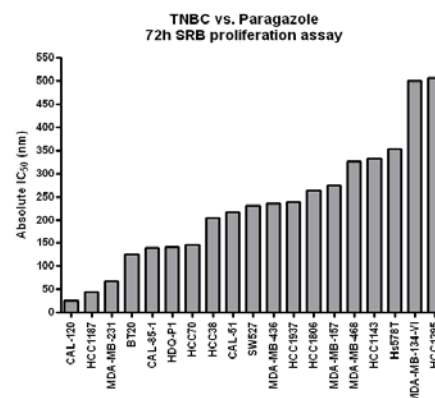


Figure 6. Growth Inhibition assays of Paragazole in TNBC lines.

Next we picked a sensitive cell line CAL-120 and a relatively growth resistance cell line Hs578t to stably over express HDAC1,3,6,8 in them using lentiviral mediate gene transfer. Of four enzymes, only overexpression of HDAC1 has observable differences in Largazole or Paragazole response. These differences are quite small and more rigorous studies are needed to establish that the effects are reproducible and statistically significant (Figure 7). Preliminary data does appear to support the hypothesis that overexpression of certain HDAC isoform may render cells more sensitivity to Largazole inhibition.

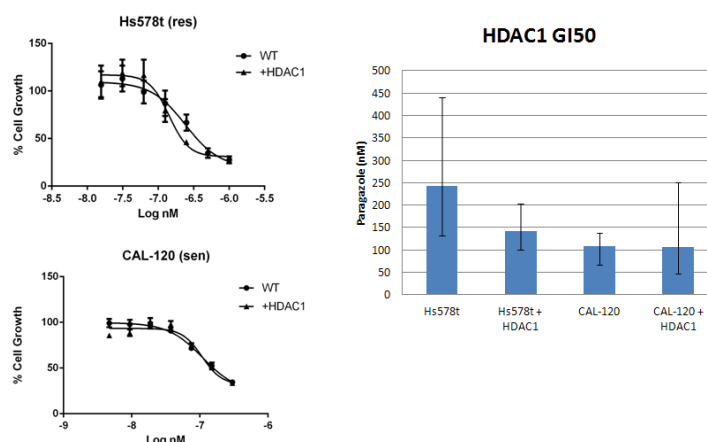


Figure 7. Growth Inhibition assays of Largazole sensitivity in cell lines overexpression of HDAC1.

Aim 3. To determine the chemotherapeutic efficacy of largazole to inhibit breast cancer growth and metastasis in mice

- Task 3.1 Large scale synthesis for largazole for initial testing along with control (Taxol) (**Phillips, Months Completed**)
- Task 3.2 Large scale synthesis of the improved lead for animal testing (**Phillips, Completed**)
- Task 3.3 Determining the chemotherapeutic effectiveness of largazole to prevent mammary tumor growth, invasion, and angiogenesis using the xenograft animal model (**Schiemann, Completed**)
- Task 3.4 Assessing the effects of largazole to prevent mammary tumor metastasis (**Schiemann, Completed**)

- Task 3.5 Testing newly improved largazole analogs in growth, invasion, angiogenesis and tumor metastasis in the xenograft animal model (**Schiemann, Months 12-23**)
- Task 3.6 Write and submit manuscripts for publication and final report to CDMRP (**Liu, Phillips and Schiemann, Month 24**)

Because Dr. William Schiemann, the collaborator on this project has moved his laboratory to Case Western Reserve University in Cleveland, it took a while for the BRCP to complete issuing the funding for this aim to us. Because of his move we have to rework the subcontract and obtain approval for IACUC and animal protocols. In addition, Professor Schiemann had to rebuild his research team at Case Western Reserve University since many experienced researchers in his lab did not follow him to Cleveland. Despite of these setbacks, Professor Schiemann group managed to make progress on this project. We made a derivative of Largazole called Paragazole. *In the last report, we mentioned that Professor Schiemann had obtained preliminary results showing that 1) Paragazole Potently Inhibits the Growth of Human and Murine TNBC Organoids (Figure 8); 2) Paragazole Potently Inhibits the Growth of Human and Murine TNBC Organoids (Figure 8). We requested a no cost extension to allow us to complete the remaining SOW. This includes repeat in vivo tumor suppression experiments. During the no cost extension period, we have shown that Paragazole (10 mg/kg) was highly effective at inhibiting the development and metastasis of human MDA-MB-231 tumors in nude mice. Importantly, the lymph node metastasis of MDA-MB-231 cells was completely absent in Paragazole treated animals (Figure 9).*

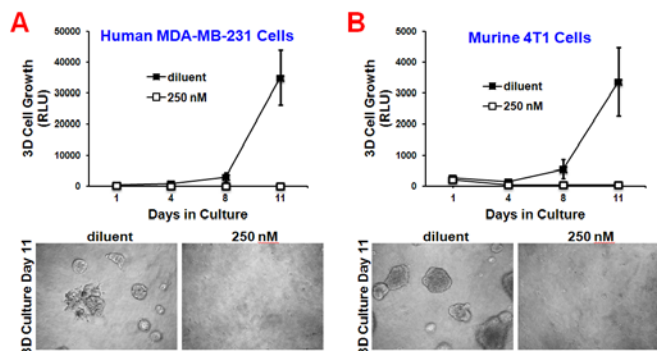


Fig. 8. Paragazole potently inhibits the growth of human and murine TNBC organoids. Human MDA-MB-231 (A) and murine 4T1 (B) TNBC cells were incubated in the absence (diluent) or presence of increasing concentrations of Paragazole (0-2 mM) over a span of 11 days. Differences in organoid growth were measured longitudinally by bioluminescence (*top*), or by bright-field microscopy (*bottom*). Data are the mean (+/-) STD of organoids treated with 250 nM of Paragazole.

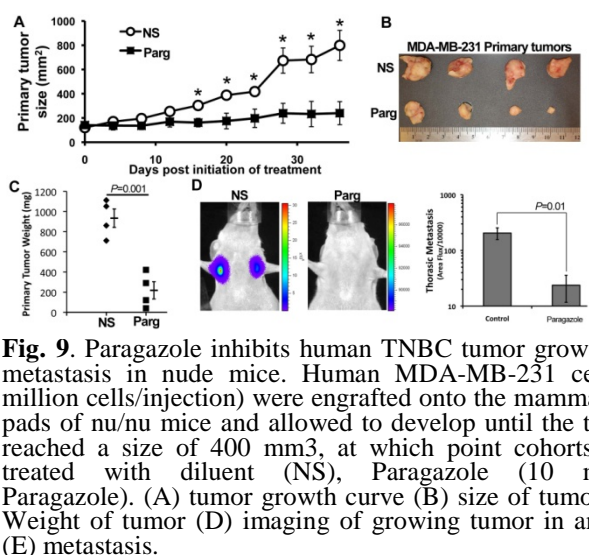


Fig. 9. Paragazole inhibits human TNBC tumor growth and metastasis in nude mice. Human MDA-MB-231 cells (2 million cells/injection) were engrafted onto the mammary fat pads of nu/nu mice and allowed to develop until the tumors reached a size of 400 mm³, at which point cohorts were treated with diluent (NS), Paragazole (10 mg/kg, Paragazole). (A) tumor growth curve (B) size of tumors (C) Weight of tumor (D) imaging of growing tumor in animals (E) metastasis.

Overall, our studies clearly demonstrated these Largazole and Paragazole are very promising in treated triple negative breast cancers in vitro and in vivo. The manuscript describing this work is current in preparation.

The Development of a novel high throughput computational tool for studying individual and collective cellular migration

Understanding how cells migrate individually and collectively during development and cancer metastasis can be significantly aided by a computation tool to accurately measure not only cellular migration speed, but also migration direction and changes in migration direction in a temporal and spatial manner. We have developed such a tool for cell migration researchers, named Pathfinder, which is capable of simultaneously measuring the migration speed, migration direction, and changes in migration directions of thousands of

cells both instantaneously and over long periods of time from fluorescence microscopy data. Additionally, we demonstrate how the Pathfinder software can be used to quantify collective cell migration. The novel capability of the Pathfinder software to measure the changes in migration direction of large populations of cells in a spatiotemporal manner will aid cellular migration research by providing a robust method for determining the mechanisms of cellular guidance during individual and collective cell migration.

3. KEY RESEARCH ACCOMPLISHMENTS

- We have completed synthesis initial round of largazole analogs for SAR testing.
- We showed that Largazole and its analogs selectively inhibit ubiquitin E1 enzyme activity in vitro
- we demonstrated that inhibitory activity of Largazole is independent of its inhibitory activity towards the histone deacetylase enzymes
- Structure-activity relationship analysis shows that the thioester bond is not required for inhibition but the macrocycle core and aliphatic tail are required.
- Largazole blocks ubiquitin activation at the adenylation step and without perturbing ubiquitin transfer from E1 to E2.
- We show that Largazole inhibition of E1 is highly selective as it does not inhibit a highly related ubiquitin E1 enzyme from *S. pombe* and is almost twenty fold less effective in inhibiting the activation of SUMO E1.
- We show Largazole represents a new class of ubiquitin E1 inhibitor.
- We show that Largazole and Paragazole potently inhibit the growth of human and murine triple negative breast cancer organoids.
- *We show Largazole and Paragazole exhibit cytotoxicity against Human MDA-MB-231 tumors in mice.*
- *We show that Pargazole inhibits triple negative breast cancer cell line Mda-mb231 metastasis to lymph nodes.*
- Overexpression of single HDAC enzyme isoform is insufficient to convert sensitive cells to resistance or vice versa.
- *Development of a novel high throughput computational tool (PathFinder) for studying individual and collective cellular migration.*
- *Using PathFinder we show that Pargazole can inhibit MDA-MB21 cell migration in response to TGF-beta.*

4. REPORTABLE OUTCOMES

Ungermannova D, Parker SJ, Nasveschuk CG, Wang W, Quade B, Zhang G, Kuchta RD, Phillips AJ, Liu X. Largazole and its derivatives selectively inhibit ubiquitin activating enzyme (e1). PLoS One. 2012;7(1):e29208.

Chapnick DA, Jacobsen J, and Liu, X. *The Development of a novel high throughput computational tool for studying individual and collective cellular migration. PLoS One. Submitted, reviewed and acceptable for publication pending minor revision.*

5. CONCLUSIONS

We have made significant progress in our proposed studies. We demonstrated that Largazole is a new class of ubiquitin E1 inhibitor and the activity of E1 inhibition is independent of its inhibitory activity toward HDAC. It is possible to design novel dual inhibitors toward both pathways. We have published one and an additional paper has been submitted and requires minor revision. We expect to publish at least two additional manuscripts on this project. We will continue to pursue the goals outlined in the original proposal in hope to develop a more effective anti-breast cancer drug candidate.

References

1. **Bowers, A., N. West, J. Taunton, S. L. Schreiber, J. E. Bradner, and R. M. Williams.** 2008. Total synthesis and biological mode of action of largazole: a potent class I histone deacetylase inhibitor. *J Am Chem Soc* **130**:11219-22.
2. **Conley, B. A., J. J. Wright, and S. Kummar.** 2006. Targeting epigenetic abnormalities with histone deacetylase inhibitors. *Cancer* **107**:832-40.
3. **Ghosh, A. K., and S. Kulkarni.** 2008. Enantioselective total synthesis of (+)-largazole, a potent inhibitor of histone deacetylase. *Org Lett* **10**:3907-3909.
4. **Haas, A. L., J. V. Warme, A. Hershko, and I. A. Rose.** 1982. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J Biol Chem* **257**:2543-8.
5. **Hershko, A., and A. Ciechanover.** 1998. The ubiquitin system. *Annu Rev Biochem* **67**:425-79.
6. **Kerscher, O., R. Felberbaum, and M. Hochstrasser.** 2006. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**:159-80.
7. **Knuesel, M., H. T. Cheung, M. Hamady, K. K. Barthel, and X. Liu.** 2005. A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. *Mol Cell Proteomics* **4**:1626-36.
8. **Minucci, S., and P. G. Pelicci.** 2006. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer* **6**:38-51.
9. **Nakayama, K. I., and K. Nakayama.** 2006. Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer* **6**:369-81.
10. **Nasveschuk, C. G., D. Ungermannova, X. Liu, and A. J. Phillips.** 2008. A concise total synthesis of largazole, solution structure, and some preliminary structure activity relationships. *Org Lett* **10**:3595-8.
11. **Nawrocki, S. T., J. S. Carew, M. S. Pino, R. A. Highshaw, R. H. Andtbacka, K. Dunner, Jr., A. Pal, W. G. Bornmann, P. J. Chiao, P. Huang, H. Xiong, J. L. Abbruzzese, and D. J. McConkey.** 2006. Aggresome disruption: a novel strategy to enhance bortezomib-induced apoptosis in pancreatic cancer cells. *Cancer Res* **66**:3773-81.
12. **Newkirk, T. L., A. A. Bowers, and R. M. Williams.** 2009. Discovery, biological activity, synthesis and potential therapeutic utility of naturally occurring histone deacetylase inhibitors. *Nat Prod Rep* **26**:1293-1320.
13. **Pagano, M., S. W. Tam, A. M. Theodoras, P. Beer-Romero, G. Del Sal, V. Chau, P. R. Yew, G. F. Draetta, and M. Rolfe.** 1995. Role of the ubiquitin-proteasome pathway in regulating abundance of the cyclin-dependent kinase inhibitor p27. *Science* **269**:682-5.
14. **Seiser, T., F. Kamena, and N. Cramer.** 2008. Synthesis and biological activity of largazole and derivatives. *Angew Chem Int Ed Engl* **47**:6483-5.
15. **Taori, K., V. J. Paul, and H. Luesch.** 2008. Structure and activity of largazole, a potent antiproliferative agent from the Floridian marine cyanobacterium *Symploca* sp. *J Am Chem Soc* **130**:1806-7.
16. **Ungermannova, D., Y. Gao, and X. Liu.** 2005. Ubiquitination of p27Kip1 requires physical interaction with cyclin E and probable phosphate recognition by SKP2. *J Biol Chem* **280**:30301-9.
17. **Wang, W., D. Ungermannova, L. Chen, and X. Liu.** 2004. Molecular and biochemical characterization of the Skp2-Cks1 binding interface. *J Biol Chem* **279**:51362-51369.
18. **Ying, Y., Y. Liu, S. R. Byeon, H. Kim, H. Luesch, and J. Hong.** 2008. Synthesis and activity of largazole analogues with linker and macrocycle modification. *Org Lett* **10**:4021-4.
19. **Ying, Y., K. Taori, H. Kim, J. Hong, and H. Luesch.** 2008. Total synthesis and molecular target of largazole, a histone deacetylase inhibitor. *J Am Chem Soc* **130**:8455-8459.
20. **Zeng, Z., W. Wang, Y. Yang, Y. Chen, X. Yang, J. A. Diehl, X. Liu, and M. Lei.** 2010. Structural basis of selective ubiquitination of TRF1 by SCFFbx4. *Dev Cell* **18**:214-25.

Largazole and Its Derivatives Selectively Inhibit Ubiquitin Activating Enzyme (E1)

Dana Ungermannova^{1‡}, Seth J. Parker^{1‡}, Christopher G. Nasveschuk^{1‡}, Wei Wang¹, Bettina Quade¹, Gan Zhang¹, Robert D. Kuchta¹, Andrew J. Phillips², Xuedong Liu^{1*}

1 Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado, United States of America, **2** Department of Chemistry, Yale University, New Haven, Connecticut, United States of America

Abstract

Protein ubiquitination plays an important role in the regulation of almost every aspect of eukaryotic cellular function; therefore, its destabilization is often observed in most human diseases and cancers. Consequently, developing inhibitors of the ubiquitination system for the treatment of cancer has been a recent area of interest. Currently, only a few classes of compounds have been discovered to inhibit the ubiquitin-activating enzyme (E1) and only one class is relatively selective in E1 inhibition in cells. We now report that Largazole and its ester and ketone analogs selectively inhibit ubiquitin conjugation to p27^{Kip1} and TRF1 *in vitro*. The inhibitory activity of these small molecules on ubiquitin conjugation has been traced to their inhibition of the ubiquitin E1 enzyme. To further dissect the mechanism of E1 inhibition, we analyzed the effects of these inhibitors on each of the two steps of E1 activation. We show that Largazole and its derivatives specifically inhibit the adenylation step of the E1 reaction while having no effect on thioester bond formation between ubiquitin and E1. E1 inhibition appears to be specific to human E1 as Largazole ketone fails to inhibit the activation of Uba1p, a homolog of E1 in *Schizosaccharomyces pombe*. Moreover, Largazole analogs do not significantly inhibit SUMO E1. Thus, Largazole and select analogs are a novel class of ubiquitin E1 inhibitors and valuable tools for studying ubiquitination *in vitro*. This class of compounds could be further developed and potentially be a useful tool in cells.

Citation: Ungermannova D, Parker SJ, Nasveschuk CG, Wang W, Quade B, et al. (2012) Largazole and Its Derivatives Selectively Inhibit Ubiquitin Activating Enzyme (E1). PLoS ONE 7(1): e29208. doi:10.1371/journal.pone.0029208

Editor: Beata G. Vertessy, Institute of Enzymology of the Hungarian Academy of Science, Hungary

Received: August 9, 2011; **Accepted:** November 22, 2011; **Published:** January 18, 2012

Copyright: © 2012 Ungermannova et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a grant from the National Institutes of Health (CA107089) to XL and United States Army grants no. W81XWH-10-1-0989 to XL and AJP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Xuedong.Liu@Colorado.edu

‡ Current address: Constellation Pharmaceuticals, Cambridge, Massachusetts, United States of America

‡ These authors contributed equally to this work.

Introduction

In humans, protein ubiquitination is a dynamic process, depending on a tightly regulated balance between the activity of two ubiquitin-activating enzymes (E1s), approximately 40 ubiquitin-conjugating enzymes (E2s), and hundreds of ubiquitin ligases (E3s). Protein ubiquitination and subsequent degradation regulates almost every aspect of eukaryotic cellular function including cell cycle regulation, endocytosis, signal transduction, apoptosis, DNA damage repair, transcriptional regulation, and many others [1]. Hershko and coworkers discovered that ubiquitin covalently modified proteins prior to their degradation in rabbit reticulocyte lysates and characterized the reaction mechanism [2]. They first described the ubiquitin-activating enzyme, E1, that carries out the ATP-dependent activation of the C-terminal glycine residue of ubiquitin prior to ligation. In the first step of the E1 activation, the enzyme forms a complex with ubiquitin and ATP and catalyzes the adenylation of ubiquitin and successive release of pyrophosphate (PP_i). During the second step, a thioester bond is formed between the C-terminus of ubiquitin and E1, subsequently releasing adenosine monophosphate (AMP). In the final step of E1 activation, additional ATP and ubiquitin are recruited to the adenylation site, generating a fully loaded E1 carrying two

molecules of ubiquitin. The activated ubiquitin is then transferred to a cysteine in the active site of ubiquitin carrier protein E2, also via thiol ester linkage. Some E2 enzymes transfer ubiquitin to acceptor proteins directly, whereas other E2s require additional substrate binding proteins known as ubiquitin ligases or E3s [3,4]. Through this mechanism, ubiquitin is attached to proteins by isopeptide linkages between the C-terminal Gly76 of ubiquitin and the ε-amino groups of lysine residues present in substrate proteins. In addition, linkages between Lys48 of one ubiquitin and the C-terminal Gly76 of another ubiquitin ultimately form polyubiquitin chains [5]. Once polyubiquitinated, proteins are targeted by the 26S proteasome for degradation.

In many human cancers, the ubiquitination system is often destabilized. For example, the cyclin-dependent kinase inhibitor p27 is mainly regulated at the protein level and is excessively degraded in approximately 50% of all human cancers [6,7]. Furthermore, expression of p27 is primarily controlled by polyubiquitination via the SCF^{Skp2} E3 ubiquitin ligase and subsequent proteasomal degradation [8]. The SCF^{Skp2} is a cullin-RING ligase (CRL), which is comprised of RING-box protein I (Rbx1), scaffold protein Cul1, linker protein Skp1, and F-box protein Skp2 [9]. In order for the ligase to function, Cul1 must first be covalently modified by NEDD8, an ubiquitin-like protein

[10–12]. Therefore, an observed stabilization of p27 in cells could result from decreased polyubiquitination by inhibiting the neddylation of Cull1 or one of the enzymes required for ubiquitination.

Given that ubiquitination influences many cellular functions, malfunctions in the pathway play a role in the pathogenesis of human neurodegenerative disorders such as Parkinson's, Alzheimer's and Huntington's diseases, as well as cancer. Inhibiting components of the ubiquitination system seems to be an avenue of therapeutic development with clinical applications [13,14]. For example, each E3 ligase targets a small number of proteins for ubiquitination, which makes it a potential target for highly specific inhibitors that have few side effects. There has, however, been little success in developing inhibitors of specific E3 ligases until recently [14,15]. Also, proteasome-inhibiting compounds have been a target of interest and were originally developed as tools for probing its proteolytic function [16,17]; however, these inhibitors were considered as possible cancer therapeutics after it was observed that they induced apoptosis in leukemic cell lines [18–20]. Although inhibiting the proteasome would nonspecifically inhibit the entire ubiquitination system, the proteasome inhibitor Bortezomib has fared surprisingly well in clinical trials and is now FDA approved for the treatment of relapsed and refractory myeloma and mantle cell lymphoma [19]. Therefore, inhibitory compounds of the ubiquitin system, whether they are specific or nonspecific, have the potential to be important therapeutics for the treatment of cancer.

In January 2008, the Leusch group at the University of Florida identified a natural product they named Largazole, which was isolated from cyanobacteria of the *Symploca* genus. They examined the compound for cytotoxicity against cancer cells and observed remarkable antiproliferative activity in transformed mammary epithelial cells. In addition, they showed that Largazole preferentially targets cancer cells over normal cells, which makes this marine substance an important synthetic target as well as a potentially valuable cancer chemotherapeutic. Remarkably, the structure consists of several unusual features, such as a 16-membered macrocycle containing a 4-methylthiazoline fused to a thiazole ring and an octanoic thioester side chain, a unit rarely found in natural products. [21]. Also, Leusch and co-workers first reported the total synthesis of Largazole and determined that the molecular basis for its anticancer activity is HDAC inhibition [21,24]. Numerous analogs of Largazole have been generated in efforts to understand the structure-activity relationship, and it has been determined that the thioester moiety is required for HDAC inhibition [21–32]. Here, we report *in vitro* mechanistic studies that reveal a potential role of Largazole as an antagonist of the ubiquitin-activating enzyme E1. In contrast to HDAC inhibition, ketone and ester analogs of Largazole can actively block the ligation of ubiquitin onto E1, indicating a differential mode of inhibitory activity since the formation of a thiol metabolite is indispensable for E1 inhibition. More explicitly, Largazole's presence negatively affected the formation of ubiquitin adenylate, which we monitored through nucleotide exchange assay.

Materials and Methods

Construction of Kip16, a GFP-p27 Expressing Cell Line

Mink lung epithelial cells expressing GFP-p27 were generated by retroviral-mediated gene transfer. pBabe-GFP-p27 amphotropic virus was made by cotransfecting pBabe-GFP-p27-Puro with pCL-Ampho in 293T cells. Viral supernatant was collected and used to infect mink lung epithelial cell line Mv 1 Lu (CCL-64) from ATCC in the presence of 8 µg/ml polybrene. Puromycin was

added at 5 µg/ml and stable clones were selected. Each clone was subcultured and tested for GFP-p27 expression in the presence or absence of 10 µM MG132 (Calbiochem, Darmstadt, Germany) for 24 hours. Clones expressing high levels of GFP in the presence of MG132 but low or undetectable GFP in its absence were expanded. Immunoblotting using an anti-p27 antibody (Santa Cruz Biotechnology, Santa Cruz, CA) was used to confirm the expression of the GFP-p27 fusion protein and stabilization of GFP-p27 upon MG132 treatment. One of the clones used for all subsequent studies was named Kip16.

Largazole Treatment of Kip 16 cells

Total synthesis of Largazole and Largazole analogs is described in [24] within the supporting information (including copies of spectra of all compounds) and is available at <http://pubs.acs.org>. Kip16 cells were seeded into 96-well flat clear-bottomed plates at 40,000 cells/well in 100 µl medium and incubated overnight at 37°C in a humidified 5% CO₂ atmosphere. Largazole was then added to final concentrations ranging from 1 µM to 1 nM in 300 µl of fresh medium. 0.3% DMSO and 1 µM of MG132 were used as negative and positive controls, respectively. After 24 hours of incubation, the medium was removed, the cells were washed twice with phosphate-buffered saline (PBS), and the cells were fixed with 4% paraformaldehyde in PBS for 15 minutes and stored at 4°C for microscopy evaluation. Cells were visualized with a GFP filter set using a 10× objective on an Eclipse TE2000-S (Nikon, Melville, NY) equipped with a Photometrics camera (Roper Scientific, Tucson, AZ).

UBA1 and His-cdc34 Purification

Human ubiquitin E1 (UBA1) was expressed with an N-terminal GST tag fusion by means of recombinant baculovirus expression in Hi5 insect cells using the pFastBacHTA vector (Invitrogen, Carlsbad, CA). The cells were lysed by sonication in the presence of protease inhibitors in a buffer containing 200 mM NaCl, 50 mM Tris-HCl pH 7.5, 1% NP40, 1 mM DTT, and 1 mM EDTA. Cleared lysate was incubated with glutathione beads (Amersham, Sweden) for one hour at 4°C. After three washes with lysis buffer, untagged E1 was produced by thrombin cleavage. The protein solution was passed through a S200 gel filtration column (Amersham, Sweden), and UBA1 concentration and purity was evaluated by SDS-PAGE and Coomassie Blue gel staining. The purity was generally greater than 90% and purified UBA1 was aliquoted and stored at –80°C after quick freezing in liquid nitrogen. N-terminal hexahistidine (His)-tagged human Cdc34 was cloned into the pQE-30 vector (Qiagen, Valencia, CA) and expressed in *Escherichia coli*. His-cdc34 was purified by Ni-NTA chromatography followed by ion exchange and size-exclusion chromatography. The purity and concentration of His-cdc34 were determined by SDS-PAGE analysis.

In Vitro Ubiquitination of p27 and Trf1

Mouse p27, cloned into pCS2, was translated *in vitro* in a reticulocyte lysate system (Promega, Madison, WI) in the presence of [³⁵S]-labeled methionine. p27 was phosphorylated by purified recombinant Cdk2-CyclinE as outlined by Ungermannova et al [33]. 5 µl of the phosphorylation reaction was incubated with a ubiquitin mixture containing 100 nM UBA1, 200 nM His-cdc34, 100 nM SCF^{Skp2} E3 ligase complex, 50 nM Cks1, 10 µM ubiquitin (Sigma Aldrich, St. Louis, MO), 10 µM methylated ubiquitin (Boston Biochem, Cambridge, MA), 1 µl of energy regeneration system (noted as 20×ER and consisting of 10 mM ATP, 20 mM Tris-HCl pH 7.4, 100 mM MgCl₂, 200 mM creatine phosphate, 2 mg/ml creatine phosphokinase, and 10%

glycerol), 1 μ M ubiquitin aldehyde, and 1 μ M MG132 in a total volume of 15 μ L. The reaction was quenched after 30 minutes in a 30°C water bath by addition of 4 \times SDS sample buffer. The products of ubiquitination were resolved by SDS-PAGE, destained in a 45% methanol and 10% acetic acid solution in water, dried and exposed overnight to a phosphorimager screen, and scanned using a Typhoon scanner 9400 (GE Healthcare, Piscataway, NJ). *In vitro* Trf1 ubiquitination was carried out as described in Zeng et al [34]. Recombinant Trf1, labeled with [γ - 33 P]-ATP by CDK1-CyclinB, was incubated with 0.5 μ M UBA1, 5 μ M UbcH5a, 1 μ M SCF^{Fbx4} E3 ligase complex, 5 μ M ubiquitin, 100 μ M methylated ubiquitin, 1 μ M ubiquitin aldehyde, and 1 μ L 20 \times ER for two hours at 30°C. Ubiquitinated Trf1 was analyzed by SDS-PAGE followed by autoradiography.

E1/E2 Thioester Bond Formation Assay

40 nM ubiquitin E1 (UBA1) or 1 μ M *S. pombe* E1 (Uba1p, gift from Chris Lima) or 0.5 μ M human SUMO E1 (Boston Biochem, Cambridge, MA) were pre-incubated with 20 \times ER at 30°C for 5 minutes in thioester reaction buffer (20 mM Tris pH 7.6, 50 mM NaCl, and 10 mM MgCl₂). After 5 minutes, 1 μ M fluorescein-ubiquitin (Boston Biochem, U-590) was added to initiate attachment of ubiquitin. All components were allowed to react for another 5 minutes in a total volume of 5 μ L. The reaction was stopped with 10 μ L of SDS-PAGE loading buffer, minus DTT, and the proteins were resolved using 12% gels that were run on ice to prevent the reduction of E1-ubiquitin due to the heat generated during electrophoresis. Thioester bond formation was visualized by scanning the gel using Typhoon scanner 9400 (GE Healthcare) that was set to fluorescence mode (532 nm). When necessary 100 nM of E2 (Cdc34) was added after the E1 enzyme was pre-charged with ATP. Serially diluted Largazole and its analogs were incubated with the reagents as stated in the text. ImageJ was utilized to quantify the fluorescence signal, and the dose response curves were generated by nonlinear least regression analysis of data using Prism (GraphPad, San Diego, CA).

[α - 32 P]-AMP: [α - 32 P]-ATP and [γ - 32 P]-PPi:[γ - 32 P]-ATP Exchange Assays

The reaction mixture contained, in a final volume of 10 μ L, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM MgCl₂ (reaction assay buffer), 150 nM human ubiquitin E1 (UBA1), 100 μ M ATP, 2 mM AMP, 1 μ M [α - 32 P]- or [γ - 32 P]-ATP (Perkin Elmer, Waltham, MA), 500 μ M PPi (sodium salt). A total of 5 μ M ubiquitin was added to the mixture to initiate the ATP:AMP exchange. After incubation at 30°C for 10 minutes, the reactions were quenched with EDTA, and 0.5 μ L aliquots of the reaction mixtures were spotted on Baker-flex[®] thin layer chromatography (TLC) polyethylenimine-modified cellulose plates (J.T. Baker, Phillipsburg, NJ) and developed in filtered 0.34 M potassium phosphate pH 7.0 for 20 minutes in a glass jar. The TLC plates were allowed to air dry for 10 minutes, covered in plastic wrap, and then exposed to a phosphorimager plate for 5–10 minutes. The separation of radiolabeled nucleotides was visualized using a Typhoon scanner 9400 (GE Healthcare, Piscataway, NJ).

Results

Largazole stabilizes GFP-p27 expression in Kip16 cells

A hallmark of many advanced cancers is an excessive degradation of the cyclin-dependent kinase inhibitor p27, which is chiefly directed by SCF^{Skp2}-mediated ubiquitination. Hence, stabilization of p27 degradation represents a rational approach in

cancer therapeutics. To identify small molecule inhibitors that can stabilize p27Kip1, we performed a screen of ~3000 compounds from NCI DTP diversity set along with several natural products in our collection. For the cell-based screen, we generated a mink lung epithelial cell line (Kip16) stably expressing p27 that was cloned in frame with green fluorescent protein (GFP). The resulting N-terminal GFP-p27 fusion, detectable by fluorescence microscopy, was used to determine the levels of p27 expression upon treatment of cells with the compound libraries in 96-well format. Much to our surprise, the most potent hit that emerged from this screen was the natural compound Largazole (L) (Figure 1), which was first described by Luesch and coworkers [21] and subsequently synthesized in several laboratories including ours [21,24–26,29,30,32]. Largazole induced a robust and highly uniform upregulation of GFP-p27 at concentrations as low as 1 nM (Figure 2A). As expected, treatment with the proteasome inhibitor MG132 is highly effective in prevention of p27 degradation. We did not observe an increase in GFP-p27 expression upon treatment with the vehicle DMSO. This result suggests that Largazole can stabilize GFP-p27 expression in cultured cells.

Largazole and select analogs inhibit the *in vitro* ubiquitination of p27 and Trf1

Before Largazole's function as an inhibitor of histone deacetylase was revealed, our initial investigation into the mechanism of this compound showed its ability to impede degradation of GFP-p27 in Kip 16 cells. One way to stabilize p27 is to block its ubiquitination. Hence we hypothesized that Largazole stabilizes p27 by inhibiting the ubiquitination pathway [7,8]. One of the downsides of cell-based assays is that the effects observed may be attributed to the influence of multiple pathways. For example, inhibiting the proteasome, elevating transcription of p27, or inhibiting Cdk activity can also lead to an increase in p27 expression. To tease out the mechanism and action of Largazole on p27 stabilization, we decided to test the effect of Largazole on p27 ubiquitination in a fully reconstituted system *in vitro* [33,35]. To test if Largazole affects p27 ubiquitination *in vitro*, we added Largazole to a p27 ubiquitin ligation reaction. As shown in Figure 2C, adding Largazole significantly reduced polyubiquitinated p27, suggesting that Largazole blocks p27 ubiquitination. Since Largazole is known to be a histone deacetylase inhibitor and has a thioester moiety that links an aliphatic chain to the core, we decided to test whether inhibition of p27 degradation can be linked to its histone deacetylase inhibitory activity. The structure-activity relationship for Largazole is relatively well understood [36]. Therefore we next tested a series of Largazole analogs (Figure 1) to study the effect of structure-activity relationship on p27 ubiquitination. To investigate this, Largazole ester (E), Largazole ketone (K), Largazole macrocycle (M), and *seco*-Largazole (S) were tested in an *in vitro* p27 ubiquitination assay (Figure 2C). We also added the HDAC inhibitor Trichostatin A (TSA), the structure of which can be found in Figure 1, to the assay to determine whether or not other HDAC inhibitors affect p27 ubiquitination. We observed that Largazole (L), Largazole ketone (K), and Largazole ester (E) inhibited the ligation of ubiquitin onto p27; however, the M and S analogs and TSA failed to inhibit the ubiquitination of p27 (Figure 2C). The fact that M had no inhibitory activity highlights the role of the octanoyl chain in hindering p27 polyubiquitination. *Seco*-Largazole (S) did not affect p27 ubiquitination, indicating the importance of the topology of the inhibitory compound. Furthermore, the result also suggests that the thioester moiety of Largazole is not required for inhibition, because the ketone and ester analogs were equally potent in blocking p27 ubiquitination. In addition, E1 inhibition is

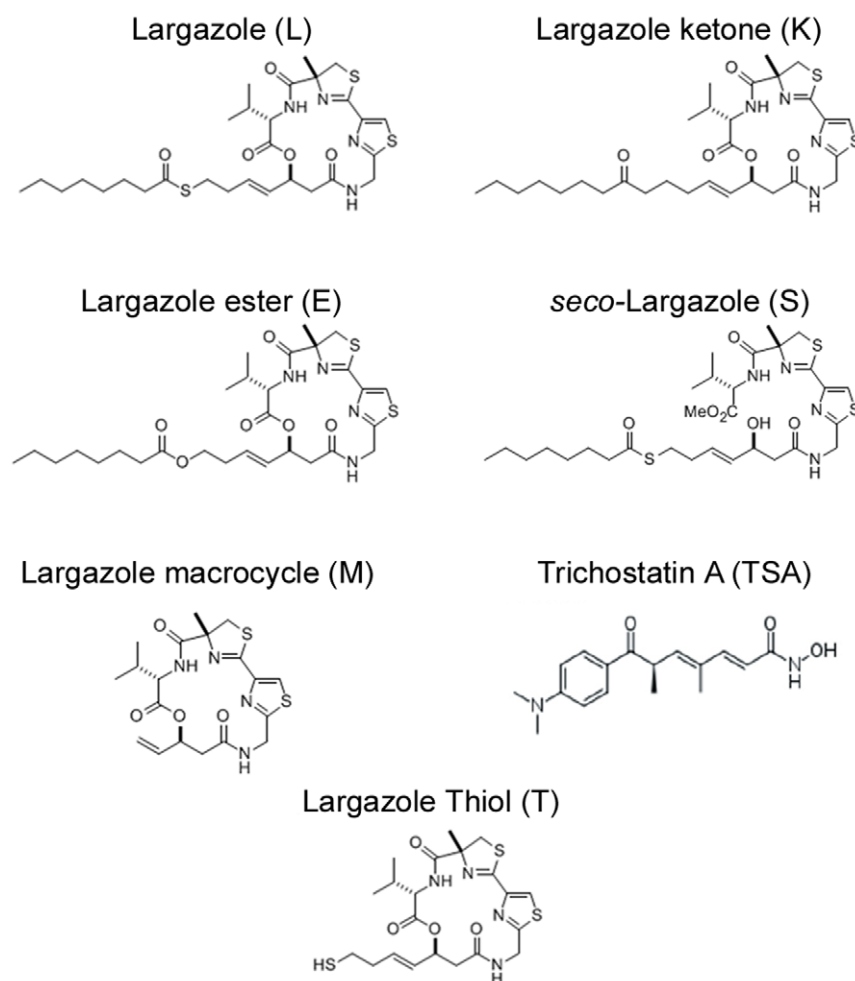


Figure 1. Chemical structures of Largazole, synthetic analogs, and Trichostatin A. Largazole (L) includes a substituted 4-methylthiazoline linearly fused to a thiazole, a 3-hydroxy-7-mercaptohept-4-enoic acid, a thioester moiety, and a hydrocarbon tail. Analogs include a substituted ketone (K) and ester (E) in place of the thioester moiety, a macrocycle lacking the thioester moiety and hydrocarbon tail (M), an analog containing a macrocycle broken at carbon-3 of the enoic acid (S), and a thiol analog lacking the thioester moiety (T). Trichostatin A (TSA) contains a hydroxamic acid functional group.

doi:10.1371/journal.pone.0029208.g001

unrelated to HDAC inhibitor activity of Largazole as both ketone and ester fail to inhibit HDAC. Prior to ubiquitination, p27 is phosphorylated by the Cdk2-CyclinE complex. We carried out an *in vitro* p27 phosphorylation assay (as described in [33]) in the presence of either DMSO or Largazole in order to test whether or not the decrease in p27 ubiquitination was due to the inhibition of the Cdk2-CyclinE complex. We observed that Largazole does not inhibit the phosphorylation of p27 (Figure 2B); therefore, the inhibition of p27-ubiquitin conjugation is due to an inhibition of the ubiquitination process rather than phosphorylation step.

To study the specificity of largazole's inhibition, we set up Trf1 *in vitro* polyubiquitination in the presence of varying concentrations of largazole ester and found that E inhibited the ubiquitin attachment in a dose-dependent manner (Figure 2D). Since ubiquitination of both proteins was impeded, and given that both reactions require different factors to execute it (p27 has to be phosphorylated by CDK2-CyclinE while there is no requirement for Trf1 phosphate addition, E2 for p27 is Cdc34, while Trf1 needs Ubc5Ha, SCFSkp2 ligase works with p27 while Fbx4 is the substrate recognition subunit for Trf1) it was evocative that largazole compounds stall a step that is common to both polyubiquitination reactions.

Largazole and ester/keto analogs inhibit ubiquitin E1 activation

Since both p27 and Trf1 can be ubiquitinated in the presence of UBA1, we hypothesized that the inhibitory activity of Largazole is due to the deactivation of E1. To test this hypothesis, we incubated Largazole and Largazole ester with recombinant E1 prior to carrying out an *in vitro* thioester assay we described previously [37]. In addition, we tested the active thiol form of Largazole (T) for E1 inhibition. The presence of a fluorescence signal in the thioester assay suggests the formation of E1-ubiquitin adducts. The dose dependent decrease in fluorescence indicates that Largazole and Largazole ester inhibit the formation of E1-ubiquitin adducts (Fig. 3AC). The dose-response curves generated (data not shown) suggest an IC_{50} of approximately 29 μ M and 25 μ M for Largazole and Largazole ester, respectively. Interestingly, the active thiol form of Largazole (T) failed to inhibit E1 (Fig. 3G), suggesting again that the octanoyl residue is important for inhibition.

Activated ubiquitin is normally transferred to ubiquitin conjugating enzymes (E2). If E1 activity is inhibited, we expect to see that defects in E1 activation should impair the attachment of

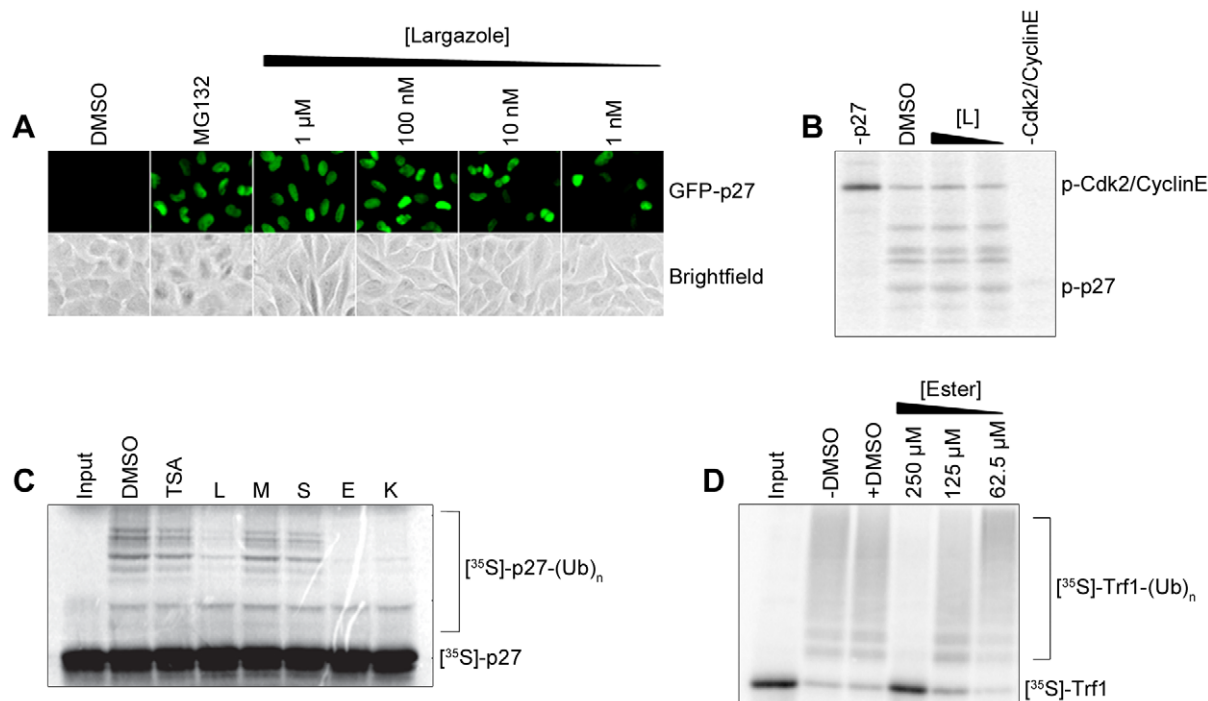


Figure 2. Largazole stabilizes p27 expression in Kip16 cells and inhibits p27 ubiquitination *in vitro*. (A) Fluorescent and corresponding bright-field images of Kip16 cells treated with varying concentrations of Largazole (L). L treatment induces the expression of GFP-p27 in a dose-dependent fashion. Addition of MG132 (1 μ M) prevents the degradation of GFP-p27 via the ubiquitination and subsequent proteasomal degradation pathway. The vehicle control, DMSO, has no effect on the reporter protein stabilization. (B) L fails to inhibit the phosphorylation of p27 by the Cdk2/CyclinE complex compared to the positive control. L (250 μ M, lane 3, and 125 μ M, lane 4) was incubated with the Cdk2/CyclinE complex prior to the autophosphorylation of Cdk2/CyclinE step. Phosphorylated-p27 was identified by protein standard. (C) L, K, and E reduce polyubiquitinated forms of p27 while M and S have no inhibitory effects. Ubiquitin-activating enzyme E1 (100 nM), UBA1, was incubated with 100 μ M of each compound prior to the reaction. (D) E reduces polyubiquitinated forms of Trf1 in a dose-dependent fashion. UBA1 (100 nM) was incubated with either DMSO or various concentrations of E ranging from 250 μ M to 1 μ M prior to the reaction. doi:10.1371/journal.pone.0029208.g002

ubiquitin onto Cdc34 (E2). To further validate E1 inhibition, we included Cdc34, the E2 enzyme required for p27 ubiquitination, in the E1 reaction mixture. As shown in Figure 3EF, in the presence of ATP, fluorescent ubiquitin is transferred to Cdc34 indicated by the presence of a fluorescent Cdc34 band on the gel. Upon incubation with E1, Largazole or Largazole ester reduce the amount of ubiquitin molecules that are transferred from E1 to E2 in a dose-dependent fashion (Fig. 3BD).

The decreased ubiquitin transfer could be attributed to either E1 or E2 inhibition; therefore, we produced E1 precharged with ubiquitin by incubating ATP and fluorescent ubiquitin for 15 minutes at room temperature followed by the addition of Cdc34, which was preincubated with either Largazole or Largazole ester. If either compound inhibits the transfer of ubiquitin from E1 to E2, then we would observe a significant decrease in Cdc34 fluorescence regardless of the order we added the compounds. Interestingly, Largazole, preincubated with Cdc34, fails to inhibit the transfer of ubiquitin from precharged E1 at concentrations <1 mM (Fig. 3E). Furthermore, in a similar experiment, Largazole ester begins to inhibit the transfer of ubiquitin from precharged E1 to Cdc34 at concentrations around 500 μ M (Fig. 3F), although this concentration is significantly above the IC_{50} of E1 inhibition. These results suggest that Largazole and Largazole ester exhibit selectivity towards ubiquitin E1. Also, this result suggests that either compound fails to promote the hydrolysis of ubiquitin thioesters on precharged E1.

Largazole ketone inhibits the adenylation step of E1 activation

E1 forms an ubiquitin-adenylate intermediate during the course of its catalytic cycle [3]. Thus the mechanism of ubiquitin E1 activation can be studied by assaying ATP:PPi and ATP:AMP exchanges [3]. Production of AMP in the $[\alpha\text{-}^{32}\text{P}]\text{-AMP}:[\alpha\text{-}^{32}\text{P}]\text{-ATP}$ exchange assay guarantees that a thioester bond is formed between E1 and ubiquitin, while the release of PPi, measured by the $[\gamma\text{-}^{32}\text{P}]\text{-PPi}:[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ exchange assay, signals the formation of ubiquitin adenylation. To further dissect the mechanism of Largazole inhibition, two nucleotide exchange assays were carried out in the presence of Largazole derivatives. For these experiments we used Largazole ketone, which is similar to Largazole and Largazole ester. From the results shown in Figure 4, it is evident that the first two concentrations of Largazole ketone (100 and 50 μ M) inhibit ubiquitination of E1 similarly and were also inhibitory in both types of exchange assays. The lack of a $[\gamma\text{-}^{32}\text{P}]\text{-PPi}$ signal suggests that the adenylation step did not occur; consequently, ubiquitin could not be transferred to the active site cysteine to trigger the release of AMP. Both steps of the E1-catalyzed reactions can be measured by the AMP:ATP exchange assay. The lack of an $[\alpha\text{-}^{32}\text{P}]\text{-AMP}$ signal further suggests that the adenylation step is inhibited by Largazole ketone. Thus Largazole or Largazole derivatives act on the first step of ubiquitin activation pathway by blocking the formation of ubiquitin-adenylate.

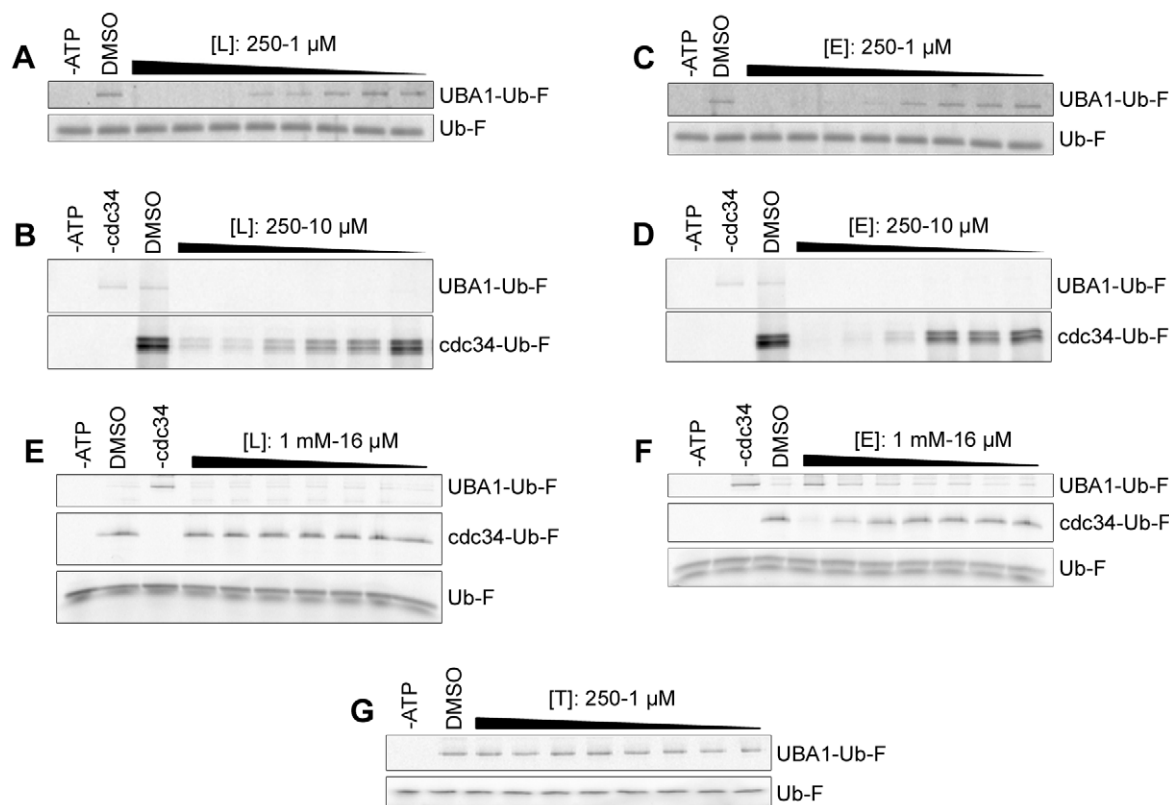


Figure 3. Largazole (L) and largazole ester (E) inhibit ubiquitin E1 in a dose dependent manner *in vitro*. (A,C) L and E inhibit transfer of ubiquitin onto E1 in a concentration-dependent manner. Thioester assay of E1 activity using fluorescein ubiquitin (Ub-F). Thioester bond formation between E1 and Ub-F is ATP-dependent (lane 2 vs. lane 1). In addition, DMSO has no effect on the formation of the thioester linkage as seen in lane 2 of both gels. 50 nM E1 was incubated with decreasing concentrations of L (A) or E (C) for 15 minutes at room temperature followed by addition of a cocktail containing ATP and Ub-F. After 5 minutes of incubation, the reactions were resolved by SDS-PAGE under non-reducing conditions. Ub-F was used to show equal loading. (B,D) Thioester assay of the ubiquitin transfer from E1 to E2 (Cdc34). Largazole or Largazole ester, when preincubated with 50 nM E1 for 15 minutes, inhibit the transfer of ubiquitin from E1 to Cdc34 in a concentration-dependent manner. (E) Largazole selectively inhibits the activity of E1 not E2. 50 nM E1 was pre-charged with ATP and then added to Cdc34 that was previously incubated with decreasing concentrations (1 mM–16 μ M) of L in thioester reaction mixture. (F) Largazole ester inhibits E2 at high concentrations. Pre-charged E1 was added to reactions that contained Cdc34 pre-incubated with E ranging from 1 mM to 16 μ M and resolved by SDS-PAGE under non-reducing conditions. Complete inhibition of ubiquitin transfer to E2 was observed at 1 mM of E, with only modest inhibition at 500 μ M. (G) Largazole thiol (T) has no effect on transfer of ubiquitin onto E1. The reaction was carried out as described in A,C.

doi:10.1371/journal.pone.0029208.g003

Selectivity of Largazole ketone against SUMO E1 and Uba1p

In addition to ubiquitin, there exist several ubiquitin-like proteins that covalently modify other proteins. All of the ubiquitin-like proteins have activation pathways similar to ubiquitin [38]. In order to study the specificity of Largazole to the ubiquitin pathway, we incubated Largazole ketone with SUMO-activating E1 enzyme prior to carrying out a thioester assay. From the results in Figure 5B, we found that Largazole ketone is ineffective in inhibiting the formation of E1-SUMO adducts. From the dose-response curve generated from the SUMO E1 fluorescence results, the IC_{50} is approximately 450 μ M as opposed to 30 μ M for ubiquitin E1 (data not shown). Thus Largazole is relatively selective in perturbing ubiquitin E1 activation.

Ubiquitin and the ubiquitin E1 enzyme are highly conserved among eukaryotes [38]. Sequence analysis shows a 45% homology between the human ubiquitin-activating enzyme E1 (UBA1) and *S. pombe* E1 (ptr3/Uba1p) at the amino acid sequence level. To test whether Largazole ketone inhibits the *S. pombe* E1, we carried out a thioester assay using Largazole ketone and the ubiquitin E1

homologue in *S. pombe*, Uba1p. The results in Figure 5A suggest that Largazole ketone fails to inhibit the formation of E1-ubiquitin adducts at concentrations less than 1 mM. Taken together, these results suggest that Largazole and its derivative are highly selective in inhibiting the ubiquitin E1 enzyme.

Discussion

In this study, we showed that Largazole and its analogs selectively inhibit ubiquitin E1 enzyme activity *in vitro*. Also, we demonstrated that the inhibitory activity of Largazole is independent of its inhibitory activity towards the histone deacetylase enzymes. Structure-activity relationship analysis shows that the thioester bond is not required for inhibition but the macrocycle core and aliphatic tail are indispensable. Largazole blocks ubiquitin activation at the adenylation step and without perturbing ubiquitin transfer from E1 to E2. Finally we show that Largazole inhibition of E1 is highly selective as it does not inhibit a highly related ubiquitin E1 enzyme from *S. pombe* and is almost twenty fold less effective in inhibiting the activation of SUMO E1. Taken together, our results reveal that Largazole represents a new class of ubiquitin E1 inhibitors.

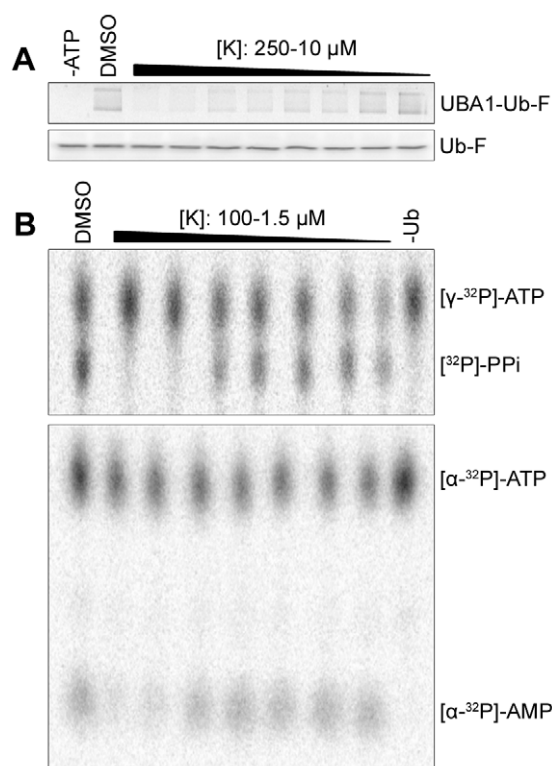


Figure 4. Largazole ketone inhibits the adenylation of the E1 ubiquitin-activating enzyme. (A) Largazole ketone inhibits ligation of ubiquitin onto E1 in a concentration-dependent fashion. Reduction of E1~Ub adducts was determined by thioester assay utilizing fluorescein ubiquitin. (B) Largazole ketone inhibits the adenylation step in ubiquitin E1 activation in a concentration-dependent fashion. K was serially diluted (100 μM to 1.5 μM) and incubated with UBA1 (150 nM) at room temperature for five minutes. The thioester reaction mixture was mixed with ubiquitin to initiate the PPi:ATP exchange (middle panel) or AMP:ATP exchange (bottom panel) and added to the UBA1/K mixture. All reactions were halted with addition of EDTA after 10 minute incubation at 37°C, resolved using Cellulose PEI TLC plates, and analyzed using a phosphorimager.
doi:10.1371/journal.pone.0029208.g004

We identified that Largazole caused a robust increase in GFP-p27 expression in Kip16 cells. This observation led us to further investigate the mechanism of GFP-p27 stabilization by Largazole. Using an *in vitro* ubiquitination assay, we were able to delineate the inhibitory point where Largazole acts on in the ubiquitination pathway, namely the E1 enzyme. However, there is a disconnect between the potency of E1 inhibition *in vitro* and GFP-p27 stabilization in cells. The EC₅₀ of Largazole for GFP-p27 stabilization is in the low nM range, yet E1 inhibition is at ~30 μM. This results suggests that the stabilization of GFP-p27 is unlikely caused by E1 inhibition, but is most likely a result of HDAC activity, which is known to block cell cycle progression and cause cell growth arrest. Consistent with this hypothesis, Largazole ketone and ester, two Largazole analogs that do not inhibit HDACs, do not increase GFP-p27 levels when Kip16 cells were treated (data not shown). However, other interpretations may account for the failure of Largazole ketone or ester to raise GFP-p27 by inhibiting E1 in cells. For example, we do not know if or how these compounds penetrate cells and how stable they are once they enter the cells. These investigations have to be undertaken before these analogs can be further developed for *in vivo* applications.

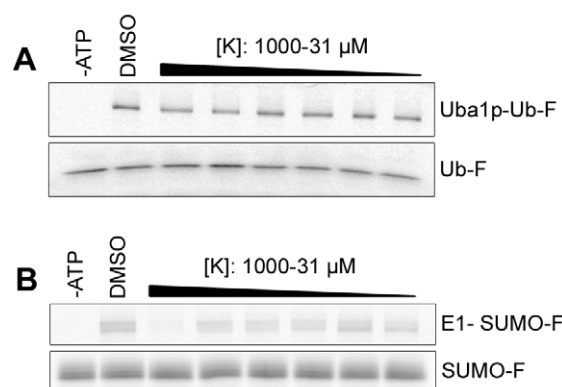


Figure 5. Investigation into the selectivity of Largazole ketone. (A) Largazole ketone (K) fails to inhibit the ligation of ubiquitin onto Uba1p, a homologue of UBA1 from *S. pombe*. Formation of Uba1p-ubiquitin adducts was determined by thioester assay utilizing fluorescein-ubiquitin. Uba1p (1.03 μM) was incubated with either DMSO or various concentrations of K serially diluted from 1000 μM to 31 μM. (B) K inhibits ligation of SUMO-1 onto human SUMO E1 in a concentration-dependent fashion. Reduction of E1-SUMO adducts was determined by thioester assay utilizing fluorescein-SUMO-1. hSUMO E1 (500 nM) was incubated with either DMSO or various concentrations of K serially diluted from 1000 μM to 31 μM.
doi:10.1371/journal.pone.0029208.g005

Panepophenanthrin, a natural compound derived from the mushroom strain *Panus rudis*, and Himeic acid A, derived from the marine fungus *Aspergillus*, are the first and second discovered inhibitors of the ubiquitin-activating enzyme E1, respectively [39,40]. Both compounds were tested *in vitro* using recombinant E1; however, the cellular activity and mechanism were not determined [41]. PYR-41 and related pyrazones are another set of compounds that were discovered to inhibit ubiquitin E1 and the first set of E1 inhibitors described to enter cells and differentially kill transformed cells [42]. The IC₅₀ of PYR-41 is around 5 μM, thus more potent than the compounds described here. However, the exact mechanism of PYR-41 inhibition is not known. Ub-AMSN represents a distinct class of protein based inhibitors of ubiquitin E1. Ub-AMSN contains a sulfamide group attached to the carboxyl terminus of ubiquitin as a nonhydrolyzable mimic of the phosphate group in the cognate Ub/Ubl-AMP adenylate intermediate. Thus, like Largazole analogs, it blocks the first step of E1 reaction [43,44]. Unfortunately, Ub-AMSN cannot be used in cells as it cannot pass through the cell membrane. However, Ub-AMSN turns out to be a very useful for probing the structure and biochemical mechanisms of E1 enzyme [44]. Therefore, Largazole and analogs could also be useful tools for probing ubiquitin function.

One of the most important questions to be answered is whether or not ubiquitin or ubiquitin-like E1 inhibitors are therapeutically relevant. Since only one ubiquitin E1 enzyme is responsible for a majority of protein ubiquitination in humans, inhibiting E1 will influence the degradation of proteins across several pathways and may lead to toxicity and, consequently, poor therapeutic efficacy. Bortezomib is the first FDA-approved proteasome inhibitor for the treatment of relapsed/refractory myeloma and mantle cell lymphoma. [19]. The proteasome, particularly the 26S proteasome, is the final step in ubiquitin-mediated protein degradation and regulates various pathways necessary for cellular function. The clinical success and efficacy of Bortezomib gives rise to the possibility that inhibitors of ubiquitin E1 will also share similar success. NEDD8 is a protein modifier that shares mechanistic and structural similarities to ubiquitin. Currently, the cullin family of

proteins has been characterized as the target for NEDD8 conjugation [11]. MLN4924 is a potent and selective inhibitor of the NEDD8-activating enzyme (NAE) that exhibited potent cytotoxicity against several human tumor-derived cell lines [45]. Interestingly, MLN4924 shares a similar mechanism to Largazole analogs. MLN4924 reacts covalently with NEDD8 mimicking a NEDD8 adenylate that is incapable of driving the reaction forward, therefore, blocking the activity of NAE [46]. MLN4924 is currently undergoing phase I clinical trials in patients with lymphoma, multiple myeloma, or any form of nonhematologic malignancies. The *in vitro* and possible clinical success of the NAE inhibitor MLN4924 further supports the concept that E1 inhibitors are potential promising cancer therapeutics.

Our preliminary structure activity relationship studies suggest that the pro-drug form of Largazole including both the hydrocarbon tail and the macrocycle are essential for E1 inhibition. For Largazole analogs to be developed as potential antitumor drugs, additional analogs are needed to be synthesized in order to improve its potency toward ubiquitin E1. The most promising aspect of Largazole analogs as ubiquitin E1 inhibitors is

the selectivity and specificity of Largazole. Largazole analogs not only display discrimination over related SUMO E1 enzyme but also remarkable selectivity in targeting human ubiquitin E1. Future structural studies would be helpful to understand how Largazole analogs inhibit E1, and insights gained from such studies may help to develop more specific inhibitors of E1. Experiments to test these hypotheses are currently underway.

Acknowledgments

We thank Eric Gunther for critical reading of the manuscript and members of Liu laboratory for helpful discussion. We thank Dr. Christopher Lima from Memorial Sloan Kettering Cancer Institute for Uba1p.

Author Contributions

Conceived and designed the experiments: DU SJP. Performed the experiments: DU SJP WW. Analyzed the data: DU SJP. Contributed reagents/materials/analysis tools: CGN BQ GZ RDK AJP. Wrote the paper: SJP DU XL.

References

- Hershko A, Ciechanover A (1982) Mechanisms of intracellular protein breakdown. *Annu Rev Biochem* 51: 335–364.
- Hershko A, Ciechanover A, Rose IA (1979) Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP. *Proc Natl Acad Sci U S A* 76: 3107–3110.
- Haas AL, Warms JV, Hershko A, Rose IA (1982) Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J Biol Chem* 257: 2543–2548.
- Hershko A, Ciechanover A, Heller H, Haas AL, Rose IA (1980) Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc Natl Acad Sci U S A* 77: 1783–1786.
- Hershko A, Heller H (1985) Occurrence of a polyubiquitin structure in ubiquitin-protein conjugates. *Biochem Biophys Res Commun* 128: 1079–1086.
- Chu IM, Hengst L, Slingerland JM (2008) The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy. *Nat Rev Cancer* 8: 253–267.
- Pagano M, Tam SW, Theodoras AM, Beer-Romero P, Del Sal G, et al. (1995) Role of the ubiquitin-proteasome pathway in regulating abundance of the cyclin-dependent kinase inhibitor p27. *Science* 269: 682–685.
- Nakayama KI, Nakayama K (2006) Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer* 6: 369–381.
- Petroski MD, Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* 6: 9–20.
- Chiba T, Tanaka K (2004) Cullin-based ubiquitin ligase and its control by NEDD8-conjugating system. *Curr Protein Pept Sci* 5: 177–184.
- Pan ZQ, Kentsis A, Dias DC, Yamoah K, Wu K (2004) Ned8 on cullin: building an expressway to protein destruction. *Oncogene* 23: 1985–1997.
- Podust VN, Brownell JE, Gladysheva TB, Luo RS, Wang C, et al. (2000) A Ned8 conjugation pathway is essential for proteolytic targeting of p27Kip1 by ubiquitination. *Proc Natl Acad Sci U S A* 97: 4579–4584.
- Nalepa G, Rolfe M, Harper JW (2006) Drug discovery in the ubiquitin-proteasome system. *Nat Rev Drug Discov* 5: 596–613.
- Hoeller D, Dikic I (2009) Targeting the ubiquitin system in cancer therapy. *Nature* 458: 438–444.
- Garber K (2005) Missing the target: ubiquitin ligase drugs stall. *J Natl Cancer Inst* 97: 166–167.
- Vinitzky A, Cardozo C, Sepp-Lorenzino L, Michaud C, Orlowski M (1994) Inhibition of the proteolytic activity of the multicatalytic proteinase complex (proteasome) by substrate-related peptidyl aldehydes. *J Biol Chem* 269: 29860–29866.
- Vinitzky A, Michaud C, Powers JC, Orlowski M (1992) Inhibition of the chymotrypsin-like activity of the pituitary multicatalytic proteinase complex. *Biochemistry* 31: 9421–9428.
- Imajoh-Ohmi S, Kawaguchi T, Sugiyama S, Tanaka K, Omura S, et al. (1995) Lactacystin, a specific inhibitor of the proteasome, induces apoptosis in human nonoblast U937 cells. *Biochem Biophys Res Commun* 217: 1070–1077.
- Orlowski RZ, Kuhn DJ (2008) Proteasome inhibitors in cancer therapy: lessons from the first decade. *Clin Cancer Res* 14: 1649–1657.
- Shinohara K, Tomioka M, Nakano H, Tone S, Ito H, et al. (1996) Apoptosis induction resulting from proteasome inhibition. *Biochem J* 317(Pt 2): 385–388.
- Taori K, Paul VJ, Luesch H (2008) Structure and activity of largazole, a potent antiproliferative agent from the Floridian marine cyanobacterium *Symploca* sp. *J Am Chem Soc* 130: 1806–1807.
- Bowers AA, West N, Newkirk TL, Troutman-Youngman AE, Schreiber SL, et al. (2009) Synthesis and histone deacetylase inhibitory activity of largazole analogs: alteration of the zinc-binding domain and macrocyclic scaffold. *Org Lett* 11: 1301–1304.
- Chen F, Gao AH, Li J, Nan FJ (2009) Synthesis and biological evaluation of c7-demethyl largazole analogues. *Chem Med Chem* 4: 1269–1272.
- Nasveschuk CG, Ungermannova D, Liu X, Phillips AJ (2008) A concise total synthesis of largazole, solution structure, and some preliminary structure activity relationships. *Org Lett* 10: 3595–3598.
- Ghosh AK, Kulkarni S (2008) Enantioselective total synthesis of (+)-largazole, a potent inhibitor of histone deacetylase. *Org Lett* 10: 3907–3909.
- Seiser T, Kamena F, Cramer N (2008) Synthesis and biological activity of largazole and derivatives. *Angew Chem Int Ed Engl* 47: 6483–6485.
- Souto JA, Vaz E, Lepore I, Poppler AC, Franci G, et al. (2010) Synthesis and biological characterization of the histone deacetylase inhibitor largazole and C7-modified analogues. *J Med Chem* 53: 4654–4667.
- Wang B, Huang PH, Chen CS, Forsyth CJ (2011) Total Syntheses of the Histone Deacetylase Inhibitors Largazole and 2-epi-Largazole: Application of N-Heterocyclic Carbene Mediated Acylations in Complex Molecule Synthesis. *J Org Chem*.
- Ying Y, Liu Y, Byeon SR, Kim H, Luesch H, et al. (2008) Synthesis and activity of largazole analogues with linker and macrocycle modification. *Org Lett* 10: 4021–4024.
- Ying Y, Taori K, Kim H, Hong J, Luesch H (2008) Total synthesis and molecular target of largazole, a histone deacetylase inhibitor. *J Am Chem Soc* 130: 8455–8459.
- Zeng X, Yin B, Hu Z, Liao C, Liu J, et al. (2010) Total synthesis and biological evaluation of largazole and derivatives with promising selectivity for cancers cells. *Org Lett* 12: 1368–1371.
- Bowers A, West N, Taunton J, Schreiber SL, Bradner JE, et al. (2008) Total synthesis and biological mode of action of largazole: a potent class I histone deacetylase inhibitor. *J Am Chem Soc* 130: 11219–11222.
- Ungermannova D, Gao Y, Liu X (2005) Ubiquitination of p27Kip1 requires physical interaction with cyclin E and probable phosphate recognition by SKP2. *J Biol Chem* 280: 30301–30309.
- Zeng Z, Wang W, Yang Y, Chen Y, Yang X, et al. (2010) Structural basis of selective ubiquitination of TRF1 by SCFFbx4. *Dev Cell* 18: 214–225.
- Wang W, Ungermannova D, Chen L, Liu X (2004) Molecular and biochemical characterization of the Skp2-Cks1 binding interface. *J Biol Chem* 279: 51362–51369.
- Newkirk TL, Bowers AA, Williams RM (2009) Discovery, biological activity, synthesis and potential therapeutic utility of naturally occurring histone deacetylase inhibitors. *Nat Prod Rep* 26: 1293–1320.
- Knuesel M, Cheung HT, Hamady M, Barthel KK, Liu X (2005) A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. *Mol Cell Proteomics* 4: 1626–1636.
- Kerscher O, Felberbaum R, Hochstrasser M (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* 22: 159–180.
- Sekizawa R, Ikano S, Nakamura H, Naganawa H, Matsui S, et al. (2002) Panepophenanthrin, from a mushroom strain, a novel inhibitor of the ubiquitin-activating enzyme. *J Nat Prod* 65: 1491–1493.

40. Tsukamoto S, Hirota H, Imachi M, Fujimuro M, Onuki H, et al. (2005) Himeic acid A: a new ubiquitin-activating enzyme inhibitor isolated from a marine-derived fungus, *Aspergillus* sp. *Bioorg Med Chem Lett* 15: 191–194.
41. Eldridge AG, O'Brien T (2010) Therapeutic strategies within the ubiquitin proteasome system. *Cell Death Differ* 17: 4–13.
42. Yang Y, Kitagaki J, Dai RM, Tsai YC, Lorick KL, et al. (2007) Inhibitors of ubiquitin-activating enzyme (E1), a new class of potential cancer therapeutics. *Cancer Res* 67: 9472–9481.
43. Olsen SK, Capili AD, Lu X, Tan DS, Lima CD (2010) Active site remodelling accompanies thioester bond formation in the SUMO E1. *Nature* 463: 906–912.
44. Lu X, Olsen SK, Capili AD, Cisar JS, Lima CD, et al. (2010) Designed semisynthetic protein inhibitors of Ub/Ubl E1 activating enzymes. *J Am Chem Soc* 132: 1748–1749.
45. Soucy TA, Smith PG, Milhollen MA, Berger AJ, Gavin JM, et al. (2009) An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* 458: 732–736.
46. Brownell JE, Sintchak MD, Gavin JM, Liao H, Bruzzese FJ, et al. (2010) Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Mol Cell* 37: 102–111.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

**The Development of A Novel High Throughput Computational Tool For Studying
Individual and Collective Cellular Migration**

Douglas A. Chapnick¹, Jeremy Jacobsen¹, and Xuedong Liu^{1,4}

¹Department of Chemistry and Biochemistry, 3415 Colorado Ave, University of Colorado,
Boulder, Colorado 80309

⁴Corresponding author:
Xuedong Liu
Tel: (303) 735-6161
Fax: (303) 735-6161
Email: liux@colorado.edu

16

17 **Abstract**

18 Understanding how cells migrate individually and collectively during development and
19 cancer metastasis can be significantly aided by a computation tool to accurately measure not
20 only cellular migration speed, but also migration direction and changes in migration direction in
21 a temporal and spatial manner. We have developed such a tool for cell migration researchers,
22 named Pathfinder, which is capable of simultaneously measuring the migration speed, migration
23 direction, and changes in migration directions of thousands of cells both instantaneously and
24 over long periods of time from fluorescence microscopy data. Additionally, we demonstrate how
25 the Pathfinder software can be used to quantify collective cell migration. The novel capability of
26 the Pathfinder software to measure the changes in migration direction of large populations of
27 cells in a spatiotemporal manner will aid cellular migration research by providing a robust
28 method for determining the mechanisms of cellular guidance during individual and collective cell
29 migration.

Introduction

Cellular migration has been shown to be an important process in cancer progression, development, tissue repair, and immune response [1-10]. As a result, a plethora of research has been performed to identify the molecular mechanisms behind how individual cells achieve migration, as well as how neighboring cells migrate cooperatively in collective migration (reviewed in [11-13] and [14], respectively). Collective migration is defined as the ability of physically interacting cells to adopt a common migration direction [14, 15]. Like individual cell migration, the collective migration of cells has been shown to be an important process in cancer progression, development and wound repair [16-23]. Such collective behavior results from each cell responding to the environmental stimuli of neighboring cells, in addition to non-cell environmental stimuli [4, 5, 14, 15, 17, 19, 20, 24-32]. Although a relatively large amount of research has been conducted to determine mechanisms behind individual cell migration, far less is known about exactly how cells migrate collectively. Furthermore, there is no standard method in the literature to quantify the 'collectiveness' behavior during collective migration [33-35].

Previous research into individual cell migration has revealed important fundamental mechanisms by which cells migrate. For instance, when an individual cell migrates on a two-dimensional (2D) surface, it projects a front end extension that can either be broad (termed a lamellipodia) or with multiple spike-like extensions (termed filipodia), which are the result of coordinated polymerization, depolymerization, and branching of the actin cytoskeleton [12, 24, 36-46]. Such coordination of actin dynamics is controlled by local recruitment of cell polarity maintain proteins, such as CDC42/Rac and Rho, which either directly or indirectly regulate actin structure, polymerization, and attachment to the extracellular matrix [37, 38, 47-51]. The attachment of the actin cytoskeleton is largely mediated by protein complexes, termed focal adhesions, which anchor the actin cytoskeleton to trans membrane integrin

receptors and the extracellular matrix [52]. The assembly of focal adhesions allows for the cell to successfully attach a front end extension to the extracellular matrix and the disassembly of focal adhesions allows a cell to detach the rear during rear end retraction [53-55]. Focal adhesion turnover and the resulting changes to the actin cytoskeleton are regulated by several kinase activities, including focal adhesion kinase (FAK), Src kinase and Rho GTPase [12, 56-60]. The temporal and spatial regulation of both the actinomyosin skeleton and focal adhesions are regulated by a complex combination of growth factor signaling and extracellular matrix protein activities, which influence the speed of actin and focal adhesion dynamics, ultimately influencing how fast a cell can migrate [61, 62].

Our current understanding of the biochemical mechanisms underlying cellular migration have been primarily the result of *in vitro* studies conducted in 2D cell culture model systems [2, 11-13, 24, 31, 32, 38, 40-43, 54, 56, 57, 60, 62-70]. However, several critical biochemical activities governing cell migration have proven to play similar roles in three dimensional (3D) model systems and *in vivo*. For instance, in 2D, 3D and *in vivo* experiments, CDC42/Rac activity determine cellular polarity [71, 72]. Similarly, FAK kinase mediates cellular migration both in 2D and 3D assays [73-75]. As a result, investigations performed in 2D assays have shed light on biochemical mechanisms that have proven to have physiological relevance. However, recent research has also revealed that there is significant differences in cell migration machinery between cells in 2D versus 3D [75-77]. Although the conclusions made in 2D migration studies will always require confirmation of physiological relevance in *in vivo* studies, they remain a valuable tool for initial investigations into the molecular mechanisms behind cellular migration compared to 3D and *in vivo* studies because they allow for tight control of experimental conditions and more accurate observation of cellular migration behavior at single cell resolution without the use of relatively complex microscopes, such as two-photon and confocal microscopes. Many of the concerns about discrepancies in biochemical mechanisms behind 2D and 3D motility may prove to be overcome by imaging

individual cell motility in 2D on soft extracellular matrices, which have been shown to be more closely similar to *in vivo* tissues than plastic or glass cell culture plates [78]. The behavior of migrating cells can be characterized by migration speed, migration direction, and migration persistence (the ability of a cell to maintain its migration direction). In 2D studies, the measurement of cell migration behavior is conducted by either manual cell tracking [79-81] or automated cell tracking [31, 32, 61, 82-85]. Such cell tracking experiments have not only shed light on how a cell achieves migration, but also have shown that cells can undergo chemotaxis towards a localized biochemical signals [20, 86, 87]. In these studies, a Dunn Chamber is used to present a chemokine gradient to cells, where cells migrate upstream of EGF and uPA gradients [88]. Such studies into how cells achieve chemotaxis highlight the need for cell migration tracking programs to not only calculate the speed and persistence of cells, but also to report the direction and changes in direction during cellular migration.

Although several computational tools exist that allow for automated cell tracking of individual cells in time-lapse microscopy videos (**Fig. 1**), these tools focus almost entirely on either the speed or persistence (ability to maintain a migration direction) or a cell. We have developed an automated high throughput cell tracking software, named Pathfinder, which is capable of simultaneously measuring and reporting cellular migration speed, migration direction and changes in migration direction of thousands of fluorescently labeled cells for an unlimited number of microscopy videos. The Pathfinder software has two improved features that distinguish it from previous cell motility computational tools, as the Pathfinder output is able to report instantaneous cell migration direction and changes in cellular migration direction. Although the cellular migration field has elucidated many fundamental mechanisms behind cellular migration, there are several key questions that remain to be answered. Specifically, 'how does a cell select a migration direction?' and 'what makes a

cell change direction?' Answering these questions will require the ability to study the migration direction and changes in migration direction at single cell and instantaneous resolution. Additionally, we explain how measurements of cellular migration direction can be used to quantify collective migration. In summary, the Pathfinder software aims to propel the cell migration field forward in both the mechanistic understanding of individual cell migration, as well as collective cell migration by allowing researchers to fully characterize cell migration behavior in an automated and high throughput manner.

Methods

Fluorescent Labeling of Cells, Cell Culture and Cellular Imaging

Stable transgenic HaCaT (Cell Lines Services, Germany) and MDA-MB-231 (ATCC, HTB-26) cell lines were fluorescently labeled via retroviral mediated gene transfer of mCherry-Histone H2B using the pRex-mCherry-H2B plasmid. For all experiments, cells were cultured in DMEM lacking phenolphthalein red and supplemented with 2 mM L-glutamine, 100 Units/mL penicillin and 100 µg/mL streptomycin. For low density assays, cells were plated at an average density of 300 cells/mm² for both HaCaT and MDA-MB-231 cells. For confluent monolayer experiments, HaCaT cells were plated at an average density of 1000 cells/mm². For epithelial sheet assays, HaCaT cells were plated at an average density of 1200 cells/mm² for 3 hours at 37 °C, after which half of partially adherent cells were manually removed using a 200 µl pipet tip. TGFβ stimulations were conducted with 100 pM ligand, while EGF stimulations were conducted with 100 nM ligand. An ImageXpress MicroXL high throughput wide-field fluorescence microscope (Molecular Devices) was used for imaging experiments at 37 °C and 5 % CO₂. All microscopy videos were acquired with a frame rate of one frame every seven minutes, in the mCherry fluorescence channel at a

magnification of 10x, where each pixel represents 1.314 μm at a pixel binning of 2 x 2. The accompanying MetaXpress software was used to compile video files from time-lapse images for each well of a 96 well plate.

Programming, Input Parameters, and Output for the Pathfinder Software

The Pathfinder software

(<https://universityofcolorado.box.com/s/qzs6nos4470sjfsaw8wg> and Fig. S1) was written in the Java programming language. User specification is required for cell radius (pixels), minimum track length, the interval of frames for desired calculations (frame n – frame n+a, where a represents the number of frames to skip for calculations), percentage of pixels in the video that represent cells, and the directory path for the folder containing .avi files (Fig. S2). The output for each video file is a single Excel spreadsheet (Fig. S3 and S4). The Pathfinder software requires only decompression of the attached .zip file and installation of JAVA runtime environment on either a 32-bit or 64-bit Windows Machine. Please note that use of pathfinder on a 64-bit machine allows for higher memory use in Java, which allows for analysis of greater numbers of cells in a single video.

Calculation of Migration Parameters, Persistence Time and Nearest Neighbor

Analyses

Cellular speed was calculated as the displacement of a cell (pixels) over 1 frame.

Conversion to $\mu\text{m}/\text{hour}$ is determined by the following equation:

$$\text{Conversion Factor} = \frac{(\text{Camera Pixel Size} \times \text{Binning Factor})}{(\text{Magnification Factor} \times \text{Frame Rate})}$$

The Angle of Trajectory was calculated from the following discontinuous equations:

$$\text{if}(dx > 0, dy > 0), \text{then } \theta_{\text{Trajectory}} = 360 - \text{ArcTAN}\left(\frac{dy}{dx}\right)$$

156 ***if***($dx > 0, dy < 0$), ***then*** $\theta_{Trajectory} = -ArcTAN\left(\frac{dy}{dx}\right)$

157 ***if***($dx < 0, dy > 0$), ***then*** $\theta_{Trajectory} = 180 - ArcTAN\left(\frac{dy}{dx}\right)$

158 ***if***($dx < 0, dy < 0$), ***then*** $\theta_{Trajectory} = 180 - ArcTAN\left(\frac{dy}{dx}\right)$

159 ***if***($dx = 0, dy = 0$), ***then*** $\theta_{Trajectory} = \text{No Angle}$

160 ***if***($dx > 0, dy = 0$), ***then*** $\theta_{Trajectory} = 0$

161 ***if***($dx < 0, dy = 0$), ***then*** $\theta_{Trajectory} = 180$

162 ***if***($dx = 0, dy < 0$), ***then*** $\theta_{Trajectory} = 90$

163 ***if***($dx = 0, dy > 0$), ***then*** $\theta_{Trajectory} = 270$

164

165 Angle of Deflection was calculated from the following discontinuous equations:

166 ***if***($\theta_{Deflection} \geq 180$), ***then*** $\theta_{Deflection} = (\theta_{Trajectory\ t=n+1} - \theta_{Trajectory\ t=n}) - 360$

167 ***if***($\theta_{Deflection} \leq -180$), ***then*** $\theta_{Deflection} = (\theta_{Trajectory\ t=n+1} - \theta_{Trajectory\ t=n}) + 360$

168

169 Persistence time calculations were performed using a modified in-house MatLab based
170 program developed by Dr. Douglas Lauffenburger (MIT) (**Fig. S5**).

171 Nearest neighbor calculations were done in Excel using an inter-centroid distance matrix
172 of (all cells) x (all cells) for each frame. Nearest neighbors were defined as cells whose
173 centroids are within 100 μm of each other. Paired random migration index of the angle of
174 trajectory (PRMI $\Theta_{Trajectory}$) calculations were also done in Excel, where the standard
175 deviation of the migration directions between pairs of neighboring cells were averaged over
176 greater than 100 pairs of cells for each condition. This average of standard deviations is
177 referred to as the PRMI $\Theta_{Trajectory}$.

178

179

Results

The overview and capabilities of the pathfinder software

The JAVA based Pathfinder software was developed to allow researchers to easily analyze large data sets of time-lapse fluorescence microscopy videos of motile cells. Since cellular tracking is already a well-established technique, our software implements a previously validated tracking algorithm ('Particle Tracker') developed by Sbalzarini *et. al* to detect each fluorescently labeled nuclei in each frame (**Fig. 2A, left**), as well as to assemble such positional information into cellular tracks (**Fig. 2A, right**), as described in their publication [89]. Since cellular positions alone are of little use to researchers in the cell migration field, we developed an analysis algorithm to transform the previous 'Particle Tracker' output into an excel spreadsheet that displays calculations of the speed, the direction, and changes in direction of individual cells, as well as the average values for a population of cells (**Fig. 2B, S3, and S4**). Simultaneous reporting of these parameters makes the Pathfinder software unique compared to other available computational motility programs (**Fig. 1**). In addition, Pathfinder is capable of running batch parallel processing of unlimited .avi files, allowing for automated and high throughput data processing of fluorescent time-lapse microscopy videos, provided they are placed in a single folder that can be navigated to from within the Pathfinder GUI. The Pathfinder program requires either a 32 bit or a 64 bit windows operating system with JAVA Runtime Environment, and is available at <https://universityofcolorado.box.com/s/qzs6nos4470sifsaw8wg> .

Fluorescent Labeling of Cells Does not Significantly Alter Cell Motility

We compared the migration speeds of wild type MDA-MB-231 cells to MDA-MB-231 cells expressing a nuclear fluorescence marker in the presence and absence of EGF in

order to determine if the introduction of nuclear marker significantly impacted ligand induced migration. Unlabeled wild type cells were manual segmented and analyzed using Pathfinder, while fluorescent images were automatically segmented and analyzed using Pathfinder. Introduction of a fluorescent nuclear marker into these cells did not significantly alter the EGF induced cellular migration speed (**Fig. S6**). However, we do not rule out that different methods of gene delivery and types of nuclear markers (for instance a fluorescent protein other than histone H2B) could lead to permanent changes in cell migration.

Using the average absolute angle of deflection to measure cellular persistence

In order to provide a means for high throughput calculation of cellular migration persistence, we used a non-traditional, but direct, approach of calculating the angle of deflection for each cell at each time. **Fig. 2B**, bottom left, illustrates how the angle of deflection measures migration persistence. The diagram represents a single cell, whose position is measured at three successive time points (1, 2, and 3, respectively). As the cell travels from 1 to 2 it maps out a line representing the trajectory of the cell between these two times. Similarly, as the cell travels from 2 to 3, another line is formed. The angle of deflection is the angle between these two lines, where a clockwise turn has a positive value, and a counterclockwise turn has a negative value. Using this calculation, each cell at each time can be assigned an angle of deflection, such that the sampling of many cells at a single time point can provide an accurate measurement of how straight cells are migrating within the population. A decrease in the average absolute value of the angle of deflection for a population of cells ($\langle |\theta_{Deflection}| \rangle$) reflects an increase in the migration persistence. We use the absolute value of the angle of deflection for describing the persistence of migration in

large populations, rather than maintaining the sign of the angle of deflection, because cells do not display a bias in which direction they prefer to turn (**Fig. 3A**).

Comparing methods to measure migration persistence

Although we measure migration persistence using the average absolute angle of deflection, the measurement of migration persistence is currently conducted in the cell migration field through the determination of persistence time, which is calculated using data fitting of time dependent mean squared displacement trends to Equation 1, where MSD is the mean squared displacement of the cell, n_d is the number of dimensions in which cells are migrating, S^2 represents the squared speed of the cell, P represents the persistence time of a cell, and t is the time.

$$MSD = n_d S^2 P t \left[1 - \frac{P}{t} (1 - e^{-tP}) \right] \quad \text{Equation 1}$$

Persistence time is used to measure migration persistence because local changes in mean squared displacements trends are likely to be associated with changes in the direction of cellular migration, provided that speed is taken into consideration. A key difference between our approach and the persistence time approach is that persistence time measurements focus on how long a cell maintains a direction, while the average absolute angle of deflection measurements focus on the degree to which cells in a population turn in each frame. Thus, persistence time measurements reflect behavior over a specified interval of time (usually 2-4 hours), while the average absolute angle of deflection measurements reflect the relatively instantaneous behavior of cells.

We compared our technique of measuring migration persistence to the method of measuring persistence time by examining the time-lapse microscopy videos of two cell lines,

MDA-MB-231 and HaCaT, stably expressing fluorescent nuclear markers and treated with no ligand, Transforming Growth Factor Beta (TGF β) or Epidermal Growth Factor (EGF). When such videos of HaCaT cells (24-26 hours post ligand stimulation) are analyzed by the two methods, both techniques lead to the same conclusions; TGF β and EGF stimulation cause increased migration persistence, where EGF has an impact of higher magnitude than that of TGF β (**Fig. 3B**). Both methods also agree when the same analysis is applied to MDA-MB-231 cells, where only TGF β has a low magnitude effect on migration persistence (**Fig. 3B**). The measured persistence time of approximately 30 minutes observed for EGF stimulated MDA-MB-231 cells is consistent with similar results from other studies [82]. Since these two techniques yield the same results under these experimental conditions, we conclude that both techniques accurately measure migration persistence in motile cells. However, there is one critical difference between our method for measuring migration persistence compared to the persistence time method. Measuring the average absolute angle of deflection can yield a measurement for migration direction for each frame, while persistence time calculations require enough frames to construct a MSD vs time plot to fit to Equation 1. Thus, the persistence time calculation, which is traditionally calculated from sampling over 2-4 hours, cannot accurately report when a cell turns and how much it turns, but instead reports its average tendency to turn. The value of having the ability to measure exactly when and to what degree a cell turns will prove useful for future investigations into the mechanism by which cells are guided and will aid researchers in answering how exactly a cell determines where to extend its front end.

Overlapping intervals suppresses noise in directional cellular behavior

The Pathfinder software is unique from other software in that it calculates angular information about individual cells. However, such angular information requires that the determination of a cell's position be relatively noise free. When we closely examine the

tracks of individual cells, we find that tracks exhibit slight vibration on the short timescale (7 minutes), such that a cell that migrates relatively straight does not display a perfectly straight track. As a result, we use overlapping intervals for our calculations of angular information in order to suppress the effects of such vibration on angular calculations. The schematic diagram in **Fig. 4A** illustrates how overlapping intervals aid in the reduction of noise in the calculation of cellular speed, direction, and persistence. Presented is the path of a single hypothetical cell that travels from positions 1 to 6. When calculating the trajectory of the movement from position 1 to 2, the resulting vector does not accurately represent the underlying trajectory of the cell over time. However, as the calculation is repeated in the same manner for a change in position from 1 to increasing successive positions, the resulting vectors quickly converge on the underlying trajectory of the cell. Each interval represents the cellular behavior in a video that has a frame rate that is the (acquisition frame rate) x (the interval size). For instance, if the interval size is 3 frames, then calculations are conducted on frames 1,4,7,10, and so forth. When successive intervals are combined, with a single frame shift from one interval to the next, the resulting data provides a time dependent parameter that has greatly suppressed noise.

Fig. 4B and C show the effect of increasing interval size on the time dependent average absolute angle of deflection trend and the average standard deviation of the absolute angle of deflection trend for HaCaT cells treated with TGF β . With an interval size of 1 frame, both the average (**Fig. 4B**) and the error (**Fig. 4C**) of the angle deflection measurements are extremely noisy. With increasing interval size, such noise is suppressed, where an interval size of greater than 2 (corresponding to 14 minutes) does not yield significant additional suppression of noise. For all cellular experiments detailed in this investigation, an interval size of 3 frames was used. This method of overlapping intervals was applied to all measurements, with the exception of persistence time measurements.

Measuring time dependent changes in migration persistence and speed

Upon mere qualitative assessment of cellular tracks, cellular behavior is difficult to deduce for large populations of cells. For example, when we examine MDA-MB-231 and HaCaT cell migration in response to either TGF β or EGF by looking at the tracks of cells between 0 and 35 hours post ligand stimulation, MDA-MB-231 cells appear to not change behavior in response to ligand treatments and HaCaT cells appear to respond in the same manner upon TGF β and EGF stimulation (**Fig. 5A and B**). However, MDA-MB-231 cells do in fact respond to ligand stimulations and HaCaT cells do in fact respond differently to TGF β and EGF stimulation (**Video S7 and S8**), which is elaborated upon below. Thus, only through rigorous quantitation can large populations be characterized for their cellular behavior.

Using our quantitative approach to measuring cellular migration, we were able to determine that either TGF β or EGF causes MDA-MB-231 cells to migrate faster, but has almost no effect on how persistently cells migrate. Only stimulation with TGF β causes a statistically significant ($p=0.003$), but extremely small in magnitude, decrease in migration persistence, as evident by a slight elevation in the average absolute angle of deflection of cells (**Fig. 5C, top left**). Both TGF β and EGF treatments yield an increase in the average speed of these cells, where EGF response is early (approximately 1 hour) and TGF β response is late (approximately 10 hours) (**Fig. 5C, bottom**). In contrast to MDA-MB-231 cells, both speed and migration persistence are activated by TGF β and EGF treatment in HaCaT cells, where the effects of EGF appear early (approximately 1 hour) and the effects of TGF β appear late (approximately 10 hours) (**Fig. 5D**). Through our rigorous quantitation method, we conclude that motility promoting ligand stimulations can differ greatly from each other in terms of both the effect on cellular migration parameters and the kinetics of

activation, both of which can be determined using the Pathfinder program. Taken together, this data illustrates the importance of time resolution in measuring cell migration responses to environmental stimuli, as well as the importance of measuring both the speed and the persistence of cells in order to characterize the full behavior of cell migration.

Quantifying collective migration with nearest neighbor analyses of angular measurements

The increasing interest in the collective behavior of cells in the cell migration field prompted us to integrate the capability of characterizing collective migration into the Pathfinder software. One already implemented method for measuring collective migration involves the calculation of the size of collectively migrating streams of cells [33]. Although this method is certainly a valid way to measure the ‘collectiveness’ of cells within a population of cells, it requires a predetermined threshold to identify whether or not neighboring cells are part of a migrating stream (a 10 degree difference in migration direction or less is required to group cells together as a collectively migrating stream). In contrast, we used Pathfinder and a matrix based calculation to provide an alternative method for measuring how similar pairs of neighboring cells are in the directions that they migrate. This method does not require any predetermined thresholds as mentioned above. For each cell, Pathfinder reports the angle of trajectory, or migration direction. For this parameter, each cell is assigned an angle that ranges from 0 to 359 degrees relative to a well-defined set of axes in the field of view. **Fig. 2B, right**, illustrates the orientation of these axes in the field of view, and provides an example of a cell that is migrating in the direction of 45 degrees. In a similar manner to the case of angle of deflection, each cell gets such an assignment for each time, providing time resolution of cellular direction. In order to demonstrate how the angle of trajectory calculation can be used to characterize collective migration, we examined HaCaT and MDA-MB-231 cells in confluent monolayers in response

to EGF stimulation. Manual inspection of the cellular tracks of these monolayers suggests that neighboring cells migrate in a similar direction in a ligand dependent manner in HaCaT cells (**Fig. 6A, top**). This behavior is not qualitatively observed for MDA-MB-231 cells (**Fig. 6A, bottom**). Since collective migration is defined as the ability of cells to adopt a common migration direction, we quantified the average standard deviation of the angle of trajectory, also referred to as the “paired random migration index” (PRMI $\Theta_{\text{Trajectory}}$), amongst pairs of nearest neighboring cells at 22-26 hours post ligand stimulation. An increase in this quantity indicates that nearest neighboring cells are migrating in increasingly different directions, meaning collective migration is decreasing. We excluded pairs of neighbors in which one cell migrates with a direction of 0-90 degrees and another cell migrates in a direction of 270-360 degrees, as these pairs would have a falsely high standard deviation due to the discontinuous transition between 360 degrees and 1 degrees. In agreement with our qualitative observations of cellular tracks, nearest neighboring HaCaT cells display a lower PRMI $\Theta_{\text{Trajectory}}$ in response to EGF stimulation (**Fig. 6B**). We compared nearest neighbor behavior to random pairing behavior in order to identify when shared behavior of cells is global versus local. If random pairing does not change the magnitude of the PRMI $\Theta_{\text{Trajectory}}$, then the shared behavior is entirely global. In contrast, when such magnitudes are affected by random pairing, the phenomenon is local. Upon random pairing of HaCaT cells, a similar trend is observed for the PRMI $\Theta_{\text{Trajectory}}$, but the magnitudes increase for both mock treatment and EGF treatment. Thus, EGF stimulation of HaCaT cells activates local collective migration, which diminishes with increasing distance between neighbors. As a control, we repeated our nearest neighbor calculations after substitution of random angles of trajectory (nearest neighbors random angles) in order to determine that the maximum value for the PRMI $\Theta_{\text{Trajectory}}$ is approximately 85 degrees in both the presence and absence of EGF stimulation. This maximum value is greater than that of the PRMI $\Theta_{\text{Trajectory}}$ in the absence of ligand stimulation, revealing that HaCaT cells do display a small degree of

collective migration, which we were not able to detect upon manual qualitative inspection of time-lapse videos. Using the same technique on MDA-MB-231 cells, we found that these cells display a statistically significant, but low magnitude, increase in the PRMI $\Theta_{\text{Trajectory}}$ amongst nearest neighbors in response to EGF, suggesting that ligand stimulation of these cells causes neighboring cells to exhibit slight repulsion, and migrate more in opposing directions upon EGF stimulation (**Fig. 6C**). Random pairing of MDA-MB-231 cells led to an increase in the magnitude of the PRMI $\Theta_{\text{Trajectory}}$, revealing that the collective migration of these cells is entirely a local phenomenon. Substitution of random angles of trajectories into data sets revealed a similar maximum value for the PRMI $\Theta_{\text{Trajectory}}$, which was approximately 85 degrees. Whether or not the apparent collective migration behavior of neighboring MDA-MB-231 cells in the absence of ligand stimulation constitutes collective migration according to the accepted definition will require further investigation into the requirement of cellular junctions in this process. However, it is worth noting that the neighboring cells do in fact have physical contact with each other (**Fig. S9**), albeit for short timescales (data not shown).

Using nearest neighbor analyses of angular measurements to characterize collective migration in epithelial sheets

In order to determine the versatility of our method of collective migration quantification, we repeated our experiments on HaCaT cells using EGF stimulation, but under the condition in which cells were arranged into epithelial sheets (**Fig. 6D and Video S10**). We asked whether or not ligand stimulation instantaneously affects all cells equally in an epithelial sheet, and found that the collective behavior response to EGF stimulation propagates away from the leading edge of an epithelial sheet overtime as indicated by propagation of decreasing PRMI $\Theta_{\text{Trajectory}}$ away from the leading edge of an epithelial sheet (**Fig. 6E**). Thus,

our technique of quantitatively characterizing collective migration has proven useful to determine the spatial collective migration properties throughout a population of cells.

Discussion

The pathfinder software provides novel useful tools for studying individual and collective cell migration

Our Pathfinder software is the first high throughput automated cell migration software to conduct measurements of instantaneous angular parameters for the characterization of cell migration behavior. Our parameter of the angle of deflection has proven useful for characterizing the migration persistence of cells in two dimensional fluorescence microscopy experiments. In addition to helping researchers characterize the molecular mechanisms behind migration persistence, this parameter can be used to identify the differential temporal responses of cells to distinct ligand stimulations, and can provide insight into the molecular mechanisms behind each ligand response. Furthermore, the available cell migration software has been entirely devoid of tools that can be used to readily characterize collective migration. As a result, investigations of collective migration have been more appropriately described as qualitative rather than quantitative. The Pathfinder software provides the first high throughput tool to quantitatively characterize collective cell migration using the angle of trajectory parameter, which can be used to determine the “Paired Random Migration Index” (PRMI $\Theta_{\text{Trajectory}}$) amongst nearest neighbors. This tool is likely to propel the collective migration field forward, as it can provide a means for accurately measuring the degree to which cells are migrating collectively, with both spatial and temporal resolution. Our method

of measuring collective migration was able to reveal hidden collective cellular behavior that cannot be easily detected qualitatively from cellular tracks, such as the local collective migration behavior amongst pairs of MDA-MB-231 cells, and the ability of EGF stimulation to suppress this behavior. In conclusion, our Pathfinder software provides novel techniques for characterizing cellular migration in a high throughput platform, which are likely to aid the cell migration field in its investigation of the molecular mechanisms behind individual and collective migration.

Acknowledgements

We would like to thank Aaron Meyer and Dr. Doug Lauffenburger for making available a MATLAB program capable of calculating migration persistence according to equation 1. We also want to thank William St. Charles, Drew Goldberg, Adam Huckaby for the initial development of the PathFinder project as a part of the required assignment for Algorithms for Molecular Biology Course taught by Professor Debra Goldberg at the University of Colorado-Boulder in 2011. We thank Dr. Golberg for her inputs and support for this project. D.A.C. was supported by a predoctoral training grant from NIGMS (T32GM08759). This work was supported by grants by US Army grant no. W81XWH-10-1-0989 and the National Institutes of Health (R01GM083172 and R01CA107098) to X.L. The ImageXpress MicroXL was supported by a NCRR grant S10 RR026680 from NIH.

Figure 1. The Pathfinder cell motility program uniquely incorporates measurements of cellular position, speed, direction, and persistence.

Although several cellular motility programs are available to measure cellular motility in terms of position and speed, only the Pathfinder program additionally reports both cellular migration direction and cellular migration persistence.

Figure 2. Angular measurements of cellular migration can reveal cellular behavior.

A) The Pathfinder program converts time-lapse microscopy videos of fluorescent HaCaT cells (left) to cellular tracks (right). **B)** The Pathfinder program uses the positional information of cellular tracks to calculate speed, where dS represents change in position and dt represents change in time (top), migration persistence through the absolute angle of deflection (bottom), and the migration direction relative to a well-defined axis orientation in the field of view (right).

Figure 3. Average absolute angle of deflection measurements accurately depict the migration persistence of cells.

A) Cells do not prefer to turn right or left in either the presence (right) or absence (left) of EGF stimulation. A binned histogram of percent of cells versus percent of right turns is normal and centered around 50 percent. **B)** A comparison of persistence time calculations and average absolute angle of deflection ($\langle |\theta_{Deflection}| \rangle$) methods for measuring migration persistence yields identical trends in both the presence and absence of either TGF β or EGF for MDA-MB-231 cells and HaCaT cells. Double asterisks indicate a p value < 0.01 . Each condition represents greater than 200 cells for persistence time measurements, and greater than 1000 cells for average absolute angle of deflection measurements.

Figure 4. Overlapping Intervals Suppress Angular Noise in Cellular Migration.

A) A schematic representation of a cellular track illustrates how increasing interval size results in the calculation of a cellular migration direction that is resistant to track vibration noise. **B)** The average absolute angle of deflection as a function of time for TGF β treated HaCaT H2B mCherry cells for an interval size of one frame shows strong scattering of measurements. Upon increasing interval size, such scattering is suppressed. **C)** Such scattering can be also measured by measuring the standard deviation in the absolute angle of deflection, which can be suppressed in a similar manner by increasing interval size. Data represents greater than 1000 cells for each plot.

Figure 5. Measuring individual cellular behavior with speed and migration persistence reveals cell type and ligand specific cellular migration behavior.

Cellular tracks of low density cells are displayed for treatments of either Mock, TGF β or EGF for MDA-MB-231 cells **(A)** or HaCaT cells **(B)**. Calibration bars represent 150 μ m. **(C)** Neither TGF- β nor EGF stimulation affects migration persistence in MDA-MB-231 cells (top). In contrast, both treatments affect cellular speed, but with different induction kinetics (bottom). **(D)** In HaCaT cells, both ligand treatments affect migration persistence and cellular speed (top and bottom, respectively). However, EGF stimulates migration persistence with earlier kinetics than that of TGF β (top), and EGF is a poor stimulator of migration speed (bottom, right). Each condition represents greater than 1000 cells.

Figure 6. Angular measurements can be used to quantify collective migration behavior.

(A) Cellular tracks of confluent monolayers of HaCaT (top) and MDA-MB-231 cells (bottom) in the presence and absence of EGF stimulation. Calibration bar represents 150 μ m. **(B)** Confluent monolayers of HaCaT cells in the presence and absence of EGF stimulation were quantified for their collective migration behavior by calculating the average standard deviation of the angle of trajectory (also called the paired random migration index (PRMI $\Theta_{\text{Trajectory}}$) amongst nearest neighboring cells. Random pairing was used to determine whether the observed behavior was local or global amongst the population. **(C)** The same quantification was conducted for MDA-MB-231 cells. **(D)** Cellular tracks of epithelial sheets of HaCaT cells in the presence and absence of EGF stimulation. Calibration bar represents 150 μ m. **(E)** Inspection of the spatial distribution of collective migration behavior reveals that EGF stimulation elicits collective migration that propagates away from the leading edge. Double asterisks indicate a p value < 0.01.

Supplementary Data S1. Pathfinder Program.

The Pathfinder program is an executable JAVA program that requires installation of JAVA Runtime Environment on a Windows operating system.

Supplementary Data S2. Parameter descriptions for the Pathfinder program GUI.

User input parameters are: Cell Outer Radius, Cell Minimum Radius (Cutoff), Percentage of Pixels with Nuclear Signal (Percentile), How Far an Average Cell is Tolerated to Migrate From Frame to Frame (Disp./Frame), How Many Frames to Bin for Calculations (Frame Binning), Minimum Tack Length (Min. Trajectory Length), Number of Parallel Threads (Threads), Folder Path for Folder with Videos (AVI Folder).

Supplementary Data S3. A description of output calculations from the Pathfinder program.

Each cell receives a cellular ID number (1), for each frame (2). In each frame a cell is assigned an X (3) and Y (4) coordinate, a displacement from the last frame in pixels (5), an angle of trajectory (6), an angle of deflection (7) and a mean squared displacement (8). Mean squared displacements can be used to calculate the persistence time for a cell. For the population of cells, Pathfinder reports the frame (9) dependent change in the average displacement (10), the average angle of trajectory (11), the percentage of cells turning greater than 90 degrees (12), and the average absolute angle of deflection (13). Additionally, Pathfinder reports a binned histogram of percent of cells versus the possible migration directions from 0 to 359 degrees (14 and 15). Lastly, the number of cellular tracks is reported (16).

Supplementary Data S4. A sample output from Pathfinder.

A sheet of HaCaT H2B-mCherry cells were analyzed for individual cell migration from a microscopy video between 22 and 25.5 hours post 100 nM EGF stimulation. Frames were acquired every 7 minutes. Analysis was conducted with a frame binning of 3 frames (21 minutes).

Supplementary Data S5. A MATLAB script to convert mean squared displacement (MSD) versus time data from a Pathfinder output into persistence time.

This MATLAB script requires user input of MSD and time data for each cell and converts this information into persistence time.

Supplementary Data S6. Wild type MDA-MB-231 cells and MDA-MB-231 H2B-mCherry cells migrate with similar speeds in the presence and absence of EGF stimulation.

Brightfield microscopy videos of mock and EGF treated wild type (WT) MDA-MB-231 cells were manually measured for position over the course of a 10 frame interval (7 minutes/frame) after 24 hours ligand or mock stimulation and the average speed of cells was calculated with a frame binning of 3. The same analysis was done on parallel videos of MDA-MB-231 using pathfinder, which yielded similar results for WT and labeled cells in the speed of migration in the presence and absence of EGF. 50 cells were used for this comparison for each condition.

Supplementary Video S7. MDA-MB-231 cells at low density upon either mock, TGF β , or EGF treatment.

MDA-MB-231 cells with an H2B-mCherry nuclear marker were observed by time-lapse microscopy using the mCherry fluorescence channel. Each frame represents 7 minutes.

Supplementary Video S8. HaCaT cells at low density upon either mock, TGF β , or EGF treatment.

HaCaT cells with an H2B-mCherry nuclear marker were observed by time-lapse microscopy using the mCherry fluorescence channel. Each frame represents 7 minutes.

Supplementary Data S9. MDA-MB-231 cells maintain physical contact with their nearest neighboring cell.

Brightfield microscopy of EGF treated MDA-MB-231 cells reveals that nearest neighboring cells have physical contact with each other.

Supplementary Video S10. Epithelial sheets of HaCaT cells upon either mock or EGF treatment.

HaCaT cells with an H2B-mCherry nuclear marker were assembled into epithelial sheets and observed by time-lapse microscopy using the mCherry fluorescence channel. Each frame represents 7 minutes.

1. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites*. Nat Rev Cancer, 2002. **2**(8): p. 563-72.
2. Friedl, P. and K. Wolf, *Tumour-cell invasion and migration: diversity and escape mechanisms*. Nat Rev Cancer, 2003. **3**(5): p. 362-74.
3. Keller, R., *Cell migration during gastrulation*. Curr Opin Cell Biol, 2005. **17**(5): p. 533-41.
4. Locascio, A. and M.A. Nieto, *Cell movements during vertebrate development: integrated tissue behaviour versus individual cell migration*. Curr Opin Genet Dev, 2001. **11**(4): p. 464-9.
5. Montell, D.J., *Border-cell migration: the race is on*. Nat Rev Mol Cell Biol, 2003. **4**(1): p. 13-24.
6. Wang, W., et al., *Identification and testing of a gene expression signature of invasive carcinoma cells within primary mammary tumors*. Cancer Res, 2004. **64**(23): p. 8585-94.
7. Wang, W., et al., *Tumor cells caught in the act of invading: their strategy for enhanced cell motility*. Trends Cell Biol, 2005. **15**(3): p. 138-45.
8. Abercrombie, M., *Contact inhibition and malignancy*. Nature, 1979. **281**(5729): p. 259-62.
9. Abercrombie, M. and J.E. Heaysman, *Observations on the social behaviour of cells in tissue culture. I. Speed of movement of chick heart fibroblasts in relation to their mutual contacts*. Exp Cell Res, 1953. **5**(1): p. 111-31.
10. Abercrombie, M. and J.E. Heaysman, *Invasiveness of sarcoma cells*. Nature, 1954. **174**(4432): p. 697-8.
11. Horwitz, R. and D. Webb, *Cell migration*. Curr Biol, 2003. **13**(19): p. R756-9.
12. Lauffenburger, D.A. and A.F. Horwitz, *Cell migration: a physically integrated molecular process*. Cell, 1996. **84**(3): p. 359-69.
13. Van Haastert, P.J. and P.N. Devreotes, *Chemotaxis: signalling the way forward*. Nat Rev Mol Cell Biol, 2004. **5**(8): p. 626-34.
14. Rorth, P., *Collective cell migration*. Annu Rev Cell Dev Biol, 2009. **25**: p. 407-29.
15. Rorth, P., *Collective guidance of collective cell migration*. Trends Cell Biol, 2007. **17**(12): p. 575-9.

16. Bronner-Fraser, M., *Neural crest cell migration in the developing embryo*. Trends Cell Biol, 1993. **3**(11): p. 392-7.
17. Lois, C., J.M. Garcia-Verdugo, and A. Alvarez-Buylla, *Chain migration of neuronal precursors*. Science, 1996. **271**(5251): p. 978-81.
18. Martin, P., *Wound healing--aiming for perfect skin regeneration*. Science, 1997. **276**(5309): p. 75-81.
19. Murase, S. and A.F. Horwitz, *Directions in cell migration along the rostral migratory stream: the pathway for migration in the brain*. Curr Top Dev Biol, 2004. **61**: p. 135-52.
20. Rorth, P., *Initiating and guiding migration: lessons from border cells*. Trends Cell Biol, 2002. **12**(7): p. 325-31.
21. Xia, Y. and M. Karin, *The control of cell motility and epithelial morphogenesis by Jun kinases*. Trends Cell Biol, 2004. **14**(2): p. 94-101.
22. Zhang, L., et al., *A role for MEK kinase 1 in TGF-beta/activin-induced epithelium movement and embryonic eyelid closure*. Embo J, 2003. **22**(17): p. 4443-54.
23. Abercrombie, M. and J.E. Heaysman, *Observations on the social behaviour of cells in tissue culture. II. Monolayering of fibroblasts*. Exp Cell Res, 1954. **6**(2): p. 293-306.
24. Ridley, A.J., et al., *Cell migration: integrating signals from front to back*. Science, 2003. **302**(5651): p. 1704-9.
25. Fischer, R.S., et al., *Local cortical tension by myosin II guides 3D endothelial cell branching*. Curr Biol, 2009. **19**(3): p. 260-5.
26. Friedl, P., et al., *Migration of coordinated cell clusters in mesenchymal and epithelial cancer explants in vitro*. Cancer Res, 1995. **55**(20): p. 4557-60.
27. Ghabrial, A., et al., *Branching morphogenesis of the Drosophila tracheal system*. Annu Rev Cell Dev Biol, 2003. **19**: p. 623-47.
28. Khalil, A.A. and P. Friedl, *Determinants of leader cells in collective cell migration*. Integr Biol (Camb), 2010. **2**(11-12): p. 568-74.
29. Prasad, M. and D.J. Montell, *Cellular and molecular mechanisms of border cell migration analyzed using time-lapse live-cell imaging*. Dev Cell, 2007. **12**(6): p. 997-1005.
30. Theveneau, E., et al., *Collective chemotaxis requires contact-dependent cell polarity*. Dev Cell, 2010. **19**(1): p. 39-53.
31. Vitorino, P., et al., *A steering model of endothelial sheet migration recapitulates monolayer integrity and directed collective migration*. Mol Cell Biol, 2011. **31**(2): p. 342-50.
32. Vitorino, P. and T. Meyer, *Modular control of endothelial sheet migration*. Genes Dev, 2008. **22**(23): p. 3268-81.
33. Slater, B., C. Londono, and A.P. McGuigan, *An algorithm to quantify correlated collective cell migration behavior*. Biotechniques, 2013. **54**(2): p. 87-92.
34. Chen, Y., et al., *Measuring collective cell movement and extracellular matrix interactions using magnetic resonance imaging*. Sci Rep, 2013. **3**: p. 1879.
35. Rosello, C., et al., *Model driven quantification of individual and collective cell migration*. Acta Biotheor, 2004. **52**(4): p. 343-63.
36. dos Remedios, C.G., et al., *Actin binding proteins: regulation of cytoskeletal microfilaments*. Physiol Rev, 2003. **83**(2): p. 433-73.
37. Cory, G.O. and A.J. Ridley, *Cell motility: braking WAVES*. Nature, 2002. **418**(6899): p. 732-3.

- 651 38. Etienne-Manneville, S. and A. Hall, *Rho GTPases in cell biology*. Nature, 2002.
652 **420**(6916): p. 629-35.
- 653 39. Itoh, R.E., et al., *Activation of rac and cdc42 video imaged by fluorescent resonance*
654 *energy transfer-based single-molecule probes in the membrane of living cells*. Mol Cell
655 Biol, 2002. **22**(18): p. 6582-91.
- 656 40. Mogilner, A. and G. Oster, *Cell motility driven by actin polymerization*. Biophys J, 1996.
657 **71**(6): p. 3030-45.
- 658 41. Ridley, A.J. and A. Hall, *The small GTP-binding protein rho regulates the assembly of*
659 *focal adhesions and actin stress fibers in response to growth factors*. Cell, 1992. **70**(3): p.
660 389-99.
- 661 42. Worthylake, R.A. and K. Burridge, *RhoA and ROCK promote migration by limiting*
662 *membrane protrusions*. J Biol Chem, 2003. **278**(15): p. 13578-84.
- 663 43. Pollard, T.D. and G.G. Borisy, *Cellular motility driven by assembly and disassembly of*
664 *actin filaments*. Cell, 2003. **112**(4): p. 453-65.
- 665 44. Welch, M.D. and R.D. Mullins, *Cellular control of actin nucleation*. Annu Rev Cell Dev
666 Biol, 2002. **18**: p. 247-88.
- 667 45. Rafelski, S.M. and J.A. Theriot, *Crawling toward a unified model of cell mobility: spatial*
668 *and temporal regulation of actin dynamics*. Annu Rev Biochem, 2004. **73**: p. 209-39.
- 669 46. Condeelis, J.S., et al., *Lamellipodia in invasion*. Semin Cancer Biol, 2001. **11**(2): p. 119-
670 28.
- 671 47. Hussain, N.K., et al., *Endocytic protein intersectin-1 regulates actin assembly via Cdc42*
672 *and N-WASP*. Nat Cell Biol, 2001. **3**(10): p. 927-32.
- 673 48. Soderling, S.H., et al., *The WRP component of the WAVE-1 complex attenuates Rac-*
674 *mediated signalling*. Nat Cell Biol, 2002. **4**(12): p. 970-5.
- 675 49. Nobes, C.D. and A. Hall, *Rho, rac, and cdc42 GTPases regulate the assembly of*
676 *multimolecular focal complexes associated with actin stress fibers, lamellipodia, and*
677 *filopodia*. Cell, 1995. **81**(1): p. 53-62.
- 678 50. Kurokawa, K., et al., *Coactivation of Rac1 and Cdc42 at lamellipodia and membrane*
679 *ruffles induced by epidermal growth factor*. Mol Biol Cell, 2004. **15**(3): p. 1003-10.
- 680 51. Moissoglu, K. and M.A. Schwartz, *Integrin signalling in directed cell migration*. Biol
681 Cell, 2006. **98**(9): p. 547-55.
- 682 52. Galbraith, C.G., K.M. Yamada, and J.A. Galbraith, *Polymerizing actin fibers position*
683 *integrins primed to probe for adhesion sites*. Science, 2007. **315**(5814): p. 992-5.
- 684 53. Burridge, K., M. Chrzanowska-Wodnicka, and C. Zhong, *Focal adhesion assembly*.
685 Trends Cell Biol, 1997. **7**(9): p. 342-7.
- 686 54. Webb, D.J., J.T. Parsons, and A.F. Horwitz, *Adhesion assembly, disassembly and*
687 *turnover in migrating cells -- over and over and over again*. Nat Cell Biol, 2002. **4**(4): p.
688 E97-100.
- 689 55. Burridge, K. and M. Chrzanowska-Wodnicka, *Focal adhesions, contractility, and*
690 *signaling*. Annu Rev Cell Dev Biol, 1996. **12**: p. 463-518.
- 691 56. Ren, X.D., et al., *Focal adhesion kinase suppresses Rho activity to promote focal*
692 *adhesion turnover*. J Cell Sci, 2000. **113** (Pt 20): p. 3673-8.
- 693 57. Webb, D.J., et al., *FAK-Src signalling through paxillin, ERK and MLCK regulates*
694 *adhesion disassembly*. Nat Cell Biol, 2004. **6**(2): p. 154-61.

58. Burridge, K., C.E. Turner, and L.H. Romer, *Tyrosine phosphorylation of paxillin and pp125FAK accompanies cell adhesion to extracellular matrix: a role in cytoskeletal assembly*. J Cell Biol, 1992. **119**(4): p. 893-903.
59. Cuevas, B.D., et al., *MEKK1 regulates calpain-dependent proteolysis of focal adhesion proteins for rear-end detachment of migrating fibroblasts*. Embo J, 2003. **22**(13): p. 3346-55.
60. Lee, J., et al., *Regulation of cell movement is mediated by stretch-activated calcium channels*. Nature, 1999. **400**(6742): p. 382-6.
61. Ware, M.F., A. Wells, and D.A. Lauffenburger, *Epidermal growth factor alters fibroblast migration speed and directional persistence reciprocally and in a matrix-dependent manner*. J Cell Sci, 1998. **111** (Pt 16): p. 2423-32.
62. Xie, H., et al., *EGF receptor regulation of cell motility: EGF induces disassembly of focal adhesions independently of the motility-associated PLCgamma signaling pathway*. J Cell Sci, 1998. **111** (Pt 5): p. 615-24.
63. Gail, M.H. and C.W. Boone, *The locomotion of mouse fibroblasts in tissue culture*. Biophys J, 1970. **10**(10): p. 980-93.
64. Hall, A., *Small GTP-binding proteins and the regulation of the actin cytoskeleton*. Annu Rev Cell Biol, 1994. **10**: p. 31-54.
65. Houk, A.R., et al., *Membrane tension maintains cell polarity by confining signals to the leading edge during neutrophil migration*. Cell, 2012. **148**(1-2): p. 175-88.
66. Mitra, S.K., D.A. Hanson, and D.D. Schlaepfer, *Focal adhesion kinase: in command and control of cell motility*. Nat Rev Mol Cell Biol, 2005. **6**(1): p. 56-68.
67. Small, J.V. and I. Kaverina, *Microtubules meet substrate adhesions to arrange cell polarity*. Curr Opin Cell Biol, 2003. **15**(1): p. 40-7.
68. Srinivasan, S., et al., *Rac and Cdc42 play distinct roles in regulating PI(3,4,5)P3 and polarity during neutrophil chemotaxis*. J Cell Biol, 2003. **160**(3): p. 375-85.
69. Weber, G.F., M.A. Bjerke, and D.W. DeSimone, *A mechanoresponsive cadherin-keratin complex directs polarized protrusive behavior and collective cell migration*. Dev Cell, 2012. **22**(1): p. 104-15.
70. Zamir, E. and B. Geiger, *Molecular complexity and dynamics of cell-matrix adhesions*. J Cell Sci, 2001. **114**(Pt 20): p. 3583-90.
71. Bray, K., C. Brakebusch, and T. Vargo-Gogola, *The Rho GTPase Cdc42 is required for primary mammary epithelial cell morphogenesis in vitro*. Small GTPases. **2**(5): p. 247-258.
72. Osmani, N., et al., *Cdc42 localization and cell polarity depend on membrane traffic*. J Cell Biol, 2010. **191**(7): p. 1261-9.
73. Sieg, D.J., et al., *FAK integrates growth-factor and integrin signals to promote cell migration*. Nat Cell Biol, 2000. **2**(5): p. 249-56.
74. Sieg, D.J., C.R. Hauck, and D.D. Schlaepfer, *Required role of focal adhesion kinase (FAK) for integrin-stimulated cell migration*. J Cell Sci, 1999. **112** (Pt 16): p. 2677-91.
75. Cukierman, E., et al., *Taking cell-matrix adhesions to the third dimension*. Science, 2001. **294**(5547): p. 1708-12.
76. Fraley, S.I., et al., *A distinctive role for focal adhesion proteins in three-dimensional cell motility*. Nat Cell Biol, 2010. **12**(6): p. 598-604.
77. Tang, H., et al., *Loss of Scar/WAVE complex promotes N-WASP- and FAK-dependent invasion*. Curr Biol, 2013. **23**(2): p. 107-17.

78. Kostic, A., C.D. Lynch, and M.P. Sheetz, *Differential matrix rigidity response in breast cancer cell lines correlates with the tissue tropism*. PLoS One, 2009. **4**(7): p. e6361.
79. Boldajipour, B., et al., *Control of chemokine-guided cell migration by ligand sequestration*. Cell, 2008. **132**(3): p. 463-73.
80. Tang, Q., et al., *Visualizing regulatory T cell control of autoimmune responses in nonobese diabetic mice*. Nat Immunol, 2006. **7**(1): p. 83-92.
81. Wolf, K., et al., *Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion*. Nat Cell Biol, 2007. **9**(8): p. 893-904.
82. Harms, B.D., et al., *Directional persistence of EGF-induced cell migration is associated with stabilization of lamellipodial protrusions*. Biophys J, 2005. **88**(2): p. 1479-88.
83. Huth, J., et al., *TimeLapseAnalyzer: multi-target analysis for live-cell imaging and time-lapse microscopy*. Comput Methods Programs Biomed. **104**(2): p. 227-34.
84. Huth, J., et al., *Significantly improved precision of cell migration analysis in time-lapse video microscopy through use of a fully automated tracking system*. BMC Cell Biol. **11**: p. 24.
85. Joslin, E.J., et al., *EGF-receptor-mediated mammary epithelial cell migration is driven by sustained ERK signaling from autocrine stimulation*. J Cell Sci, 2007. **120**(Pt 20): p. 3688-99.
86. Platek, A., et al., *v-Src accelerates spontaneous motility via phosphoinositide 3-kinase, phospholipase C and phospholipase D, but abrogates chemotaxis in Rat-1 and MDCK cells*. J Cell Sci, 2004. **117**(Pt 20): p. 4849-61.
87. Sawyer, C., et al., *Regulation of breast cancer cell chemotaxis by the phosphoinositide 3-kinase p110delta*. Cancer Res, 2003. **63**(7): p. 1667-75.
88. Sturge, J., J. Hamelin, and G.E. Jones, *N-WASP activation by a beta1-integrin-dependent mechanism supports PI3K-independent chemotaxis stimulated by urokinase-type plasminogen activator*. J Cell Sci, 2002. **115**(Pt 4): p. 699-711.
89. Sbalzarini, I.F. and P. Koumoutsakos, *Feature point tracking and trajectory analysis for video imaging in cell biology*. J Struct Biol, 2005. **151**(2): p. 182-95.

Figure 1
[Click here to download high resolution image](#)

Software Name	Website URL	Position	Speed	Direction	Persistence
ImarisTrack	http://www.bitplane.com/go/products/imaristrack	✓	✓		
Volocity	http://www.perkinelmer.com/pages/020/cellularimaging/products/volocity.xhtml	✓	✓		
Cell Track	http://db.cse.ohio-state.edu/CellTrack/	✓	✓		
Cell Tracker	http://www2.warwick.ac.uk/fac/sci/systemsbiology/staff/bretschneider/celltracker	✓	✓		
Mtrack2	http://user.interface.org.nz/~gringer/hacking/mtrack3.html	✓	✓		
MetaMorph	http://www.bioimaging-solutions.com/MolDev/metamorph_analysis.html	✓	✓		
Cell Motility Monitoring Project	http://genesis.ugent.be/cell_motility/distrib/distribution.html		✓		✓
Chemotaxis and Migration Tool	http://ibidi.com/xtproducts/en/S-Software-and-Image-Analysis/Manual-Image-Analysis/Chemotaxis-and-Migration-Tool		✓	✓	
Pathfinder	https://universityofcolorado.box.com/s/qzs6nos4470sjfsaw8wg	✓	✓	✓	✓

Figure 2
[Click here to download high resolution image](#)

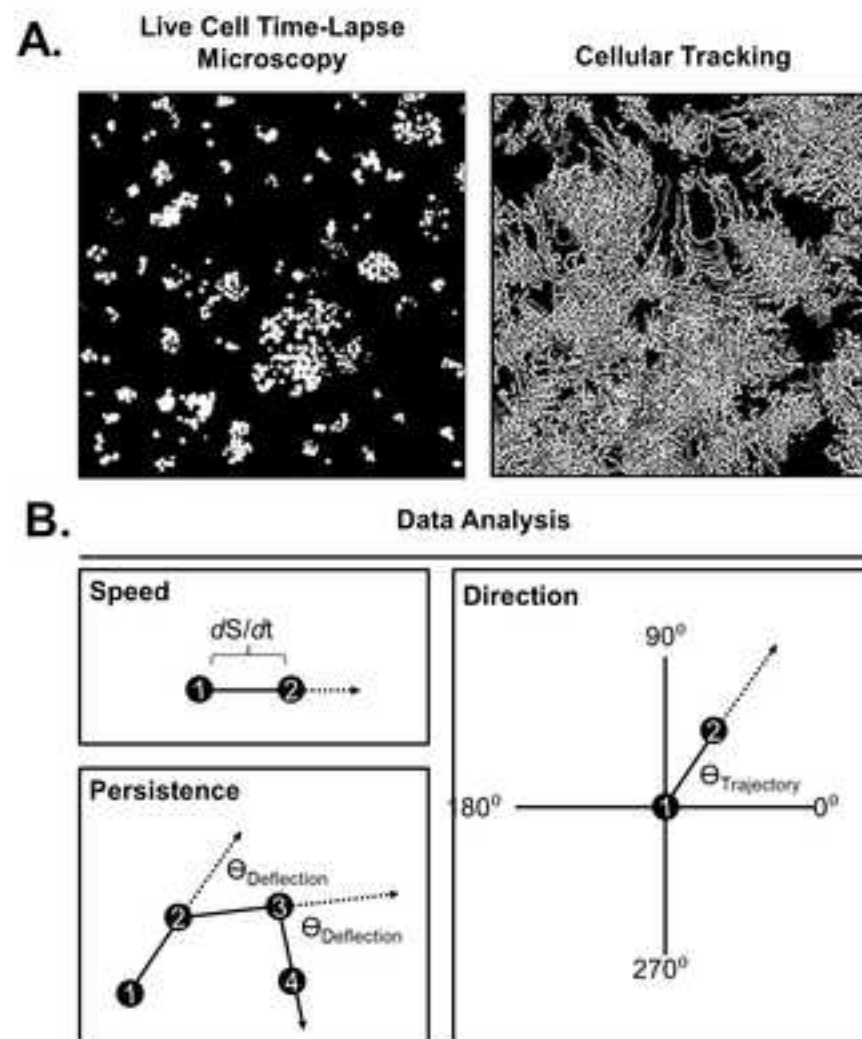


Figure 3
[Click here to download high resolution image](#)

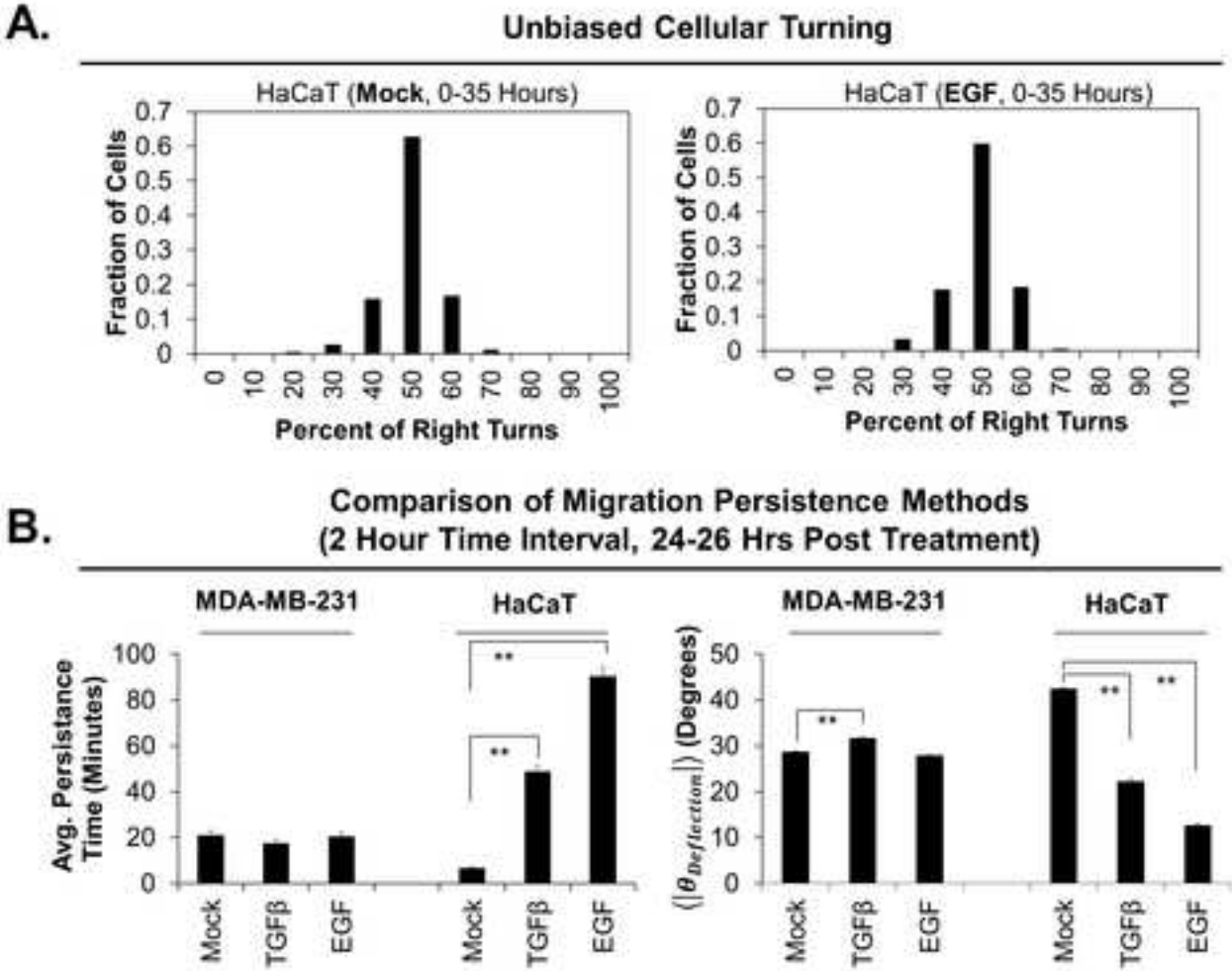


Figure 4
[Click here to download high resolution image](#)

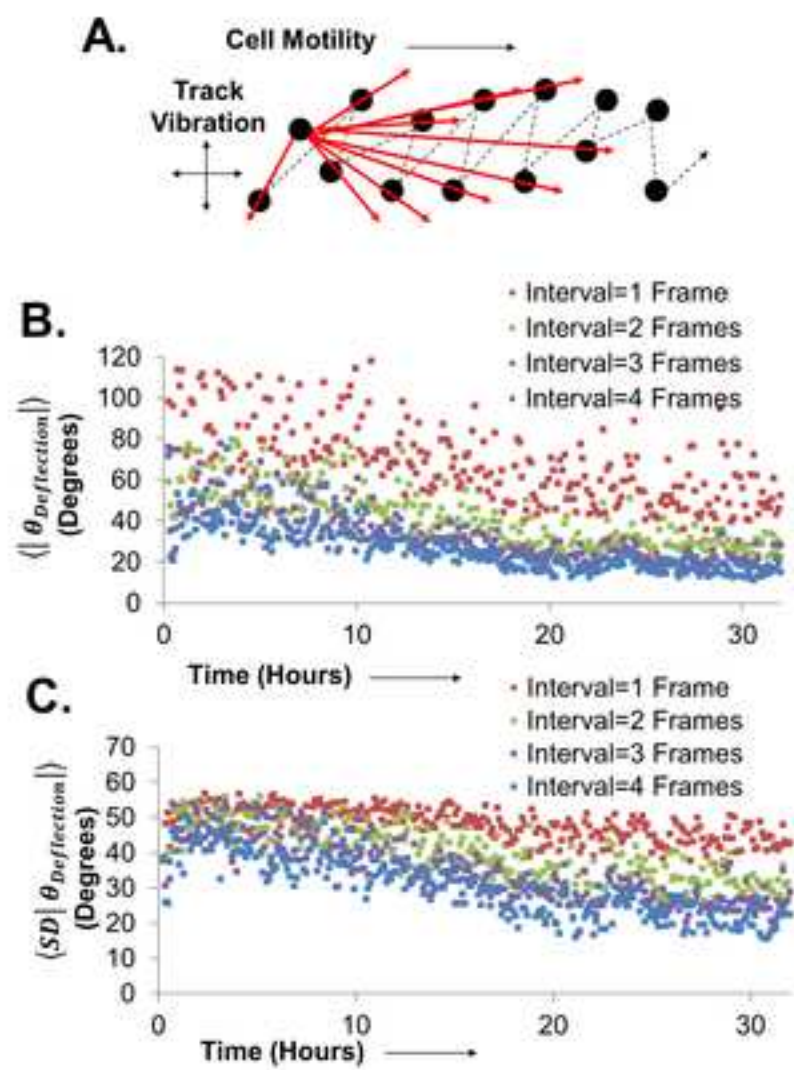


Figure 5
[Click here to download high resolution image](#)

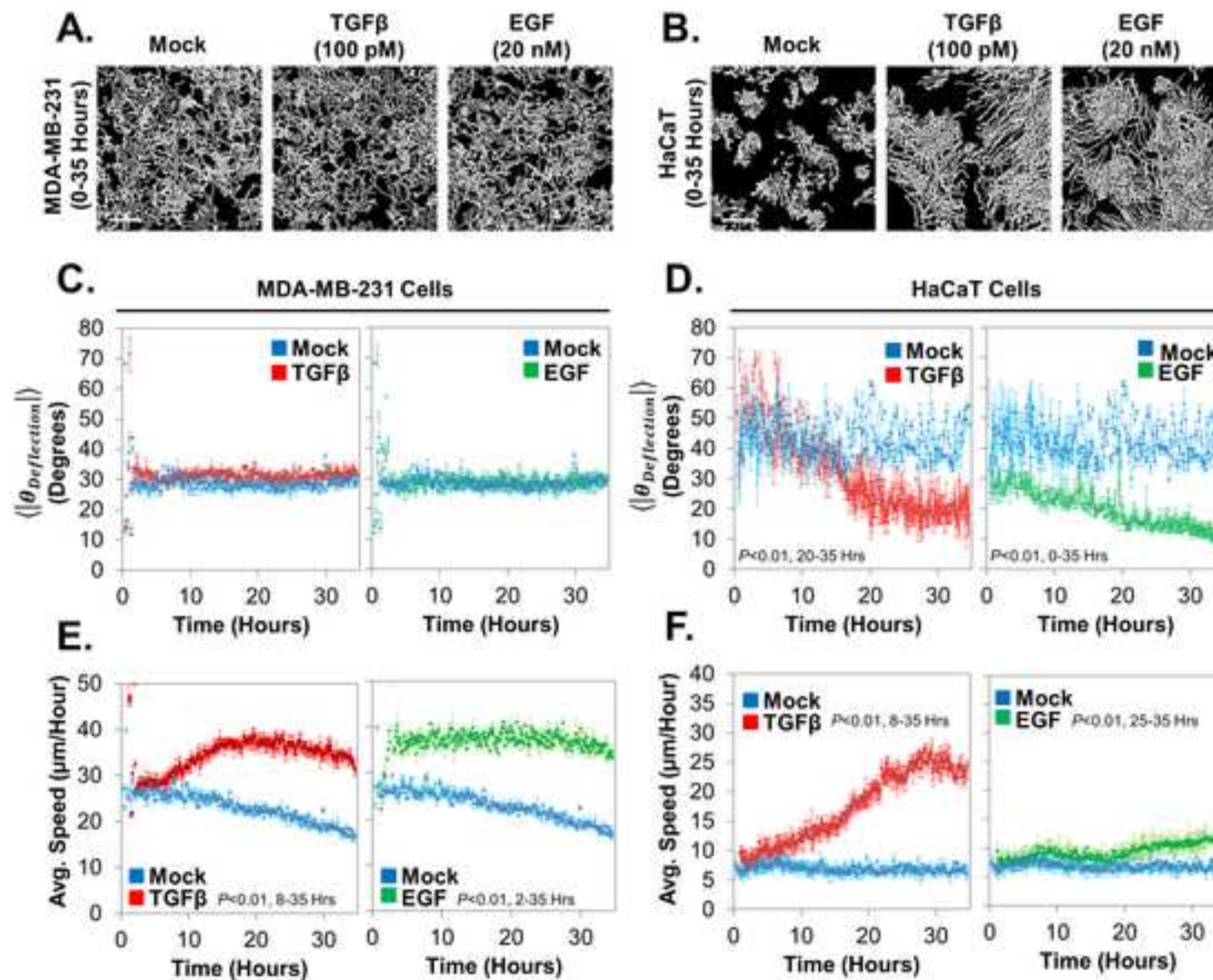


Figure 6
[Click here to download high resolution image](#)

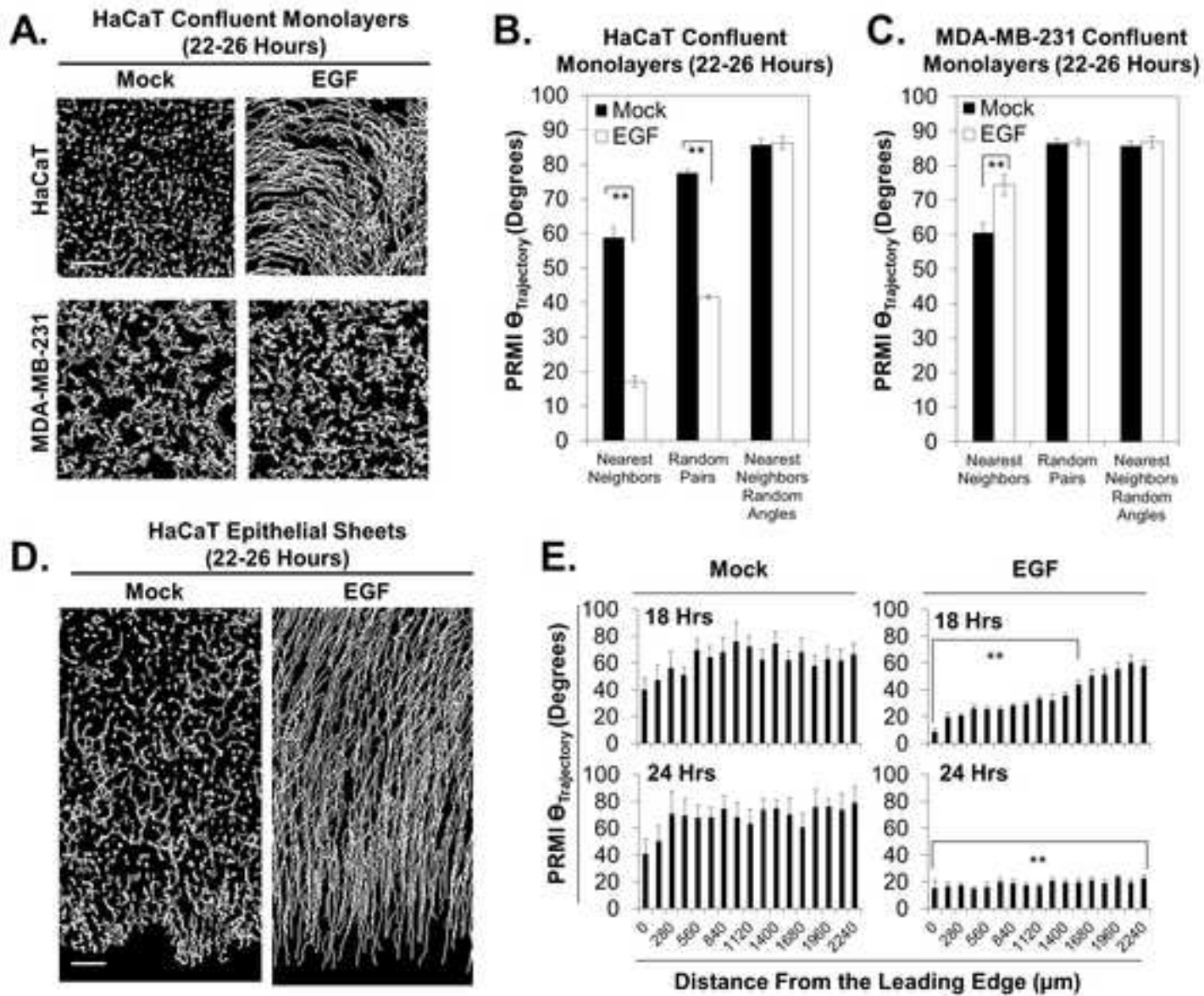


Figure S2

[Click here to download Supporting Information: Figure S2.tif](#)

Figure S3

[Click here to download Supporting Information: Figure S3.tif](#)

Figure S4

[Click here to download Supporting Information: Supplementary Data S4.xlsx](#)

Figure S5

[Click here to download Supporting Information: Supplementary Data S5.txt](#)

Figure S6

[Click here to download Supporting Information: Figure S6.tif](#)

Figure S7

[Click here to download Supporting Information: Supplementary Video S7.avi](#)

Figure S8

[Click here to download Supporting Information: Supplementary Video S8.avi](#)

Figure S9

[Click here to download Supporting Information: Figure S9.tif](#)

Figure S10

[Click here to download Supporting Information: Supplementary Video S10.avi](#)

Figure S1

[Click here to download Supporting Information - Compressed/ZIP File Archive: PathFinder64__5__31.jar](#)